ÉCOLE NORMALE SUPÉRIEURE DU BURUNDI

DÉPARTEMENT DES SCIENCES NATURELLES

SECTION : MATHÉMATIQUE

PROBABILITÉS ET STATISTIQUE ANALYTIQUE

CODE : MAT2412

UE1 : MATHÉMATIQUES APPLIQUÉES III

VOLUME : 60H (4 ECTS)

COURS MAGISTRAL (CM) : 45H

TRAVAUX DIRIGÉS (TD) : 15H

TRAVAUX PRATIQUES (TP) : 0H

SYLLABUS DE L'ÉLÉMENT CONSTITUTIF DE L'UNITÉ D'ENSEIGNEMENT (ECUE) DESTINÉ AUX ÉTUDIANTS DE BACCALAURÉAT 2 EN MATHÉMATIQUES

Titulaire : Prof. Emmanuel BARANKANIRA

Docteur en Biostatistique de l'Université de Montpellier (France, 2016)

Master en Mathématiques-Biostatistique de SupAgro (France, 2012)

DEA en Statistique, parcours Épidémiologie et Biostatistique de l'UCL (Belgique, 2008)

Licencié en Pédagogie Appliquée, Agrégé de l'Enseignement en Maths (UB, 2004)

Bujumbura, 10 octobre 2024

Descriptif du cours

Processus	Paramètres	Description
Élaboration	Titre de l'ECUE	**Probabilités et statistique analytique**
	Objectif général	Amener l'étudiant à être capable de calculer les probabilités et d'utiliser adéquatement des tests statistiques pour résoudre des problèmes de société.
	Objectifs spécifiques	À la fin de l'ECUE, l'étudiant devrait être capable de : - Calculer les probabilités des événements ; - Calculer les moments d'une variable aléatoire ; - Construire des intervalles de confiance des paramètres de population ; - Résoudre un problème qui fait appel à un test statistique.
	Prérequis	Statistique descriptive
	Organisation de l'ECUE	VHP=60h, CM=45h, TD=15h,TP=0h, TPE=40h, TGE=100h (4 crédits)
	Bref contenu du cours	Cet ECUE comprend deux parties : la partie Probabilités et la partie Statistique analytique. Avant d'aborder le calcul des probabilités, on introduit les notions d'analyse combinatoire (combinaison, arrangement, permutation, disposition ordonnée, disposition non ordonnée) dans le cas où il y a répétitions et dans le cas où il n'y en a pas. Chaque notion est accompagnée de formules et d'exercices. Ces formules classiques sont démontrées et les équations en analyse combinatoire résolues. Le calcul des probabilités commence par les définitions d'une probabilité (formule classique, utilisation de la limite, axiomatique de Kolmogorov), le modèle probabiliste, les exemples d'expériences aléatoires (univers, cardinal), les événements (simples, composés),

		la tribu borélienne, s -algèbre, l'espace probabilisable et l'espace probabilisé. Les notations ensemblistes sont comparées aux notations probabilistes. Il vient ensuite le calcul des probabilités, le calcul des probabilités conditionnelles (théorème de Bayes), l'indépendance en probabilité, la règle de multiplication, l'incompatibilité des événements, la loi jointe, la loi marginale. La notion de variable aléatoire (variable continue, variable discrète) et ses moments (espérance mathématique et variance) sont abordés. Cela conduit aux lois discrètes (loi de Dirac, loi uniforme discrète, loi de Bernoulli, loi binomiale, géométrique, hypergéométrique, loi de Poisson, loi binomiale négative) et aux lois continues (loi uniforme continue, loi exponentielle, loi normale, loi normale standard, loi log-normale, loi de Student, loi du chi-deux, loi de Fisher-Snedecor) et leurs moments. En statistique analytique, on aborde les notions de théorie d'estimation et de distribution d'échantillonnage, les techniques d'échantillonnage, l'estimation par intervalle de confiance et les tests d'hypothèse (tests de Student, test de Welch, test de Fisher, tests du chi-deux).
Méthodologie et supports pédagogiques	Méthodologie	Méthode expositive et participative
	Supports	Syllabus de l'ECUE
Modes d'évaluation	Évaluation formative	Travaux dirigés et devoirs : 40 %
	Évaluation sommative	Examen final écrit : 60 %

Table des matières

Liste des tableaux

Liste des figures

Introduction

Statistique

La statistique n'est pas une discipline récente. Son histoire date des années 1749. Étymologiquement parlant, la statistique vient du mot latin « *status* » qui signifie *État*. Dans ce sens, elle concernait les affaires relatives à l'État telles que la composition de la population, les besoins de l'armée, les impôts payés par les citoyens, les productions agricoles et les maladies contagieuses.

Actuellement, la statistique s'est étendue à tous les domaines de la vie du pays : économie, environnement, santé, agriculture, etc. La statistique se veut être une discipline qui permet de collecter des données relatives à un phénomène donné, de les traiter, de les analyser, d'en interpréter les résultats et de présenter ces derniers afin que les données soient compréhensibles par tous. La statistique est une science qui a pour objet de collecter, de rassembler et de synthétiser numériquement et/ou graphiquement l'information contenue dans les données afin de prendre des décisions, ce qui est une ébauche de définition. Elle est considérée comme étant à la fois une science, une méthode et un ensemble de techniques.

La statistique est composée principalement de deux parties : la statistique descriptive et la statistique analytique. Cette dernière partie englobe la statistique inférentielle et la statistique mathématique. La statistique descriptive, appelée également analyse exploratoire des données, a pour but de rassembler, d'ordonner, de classer, de synthétiser et de représenter les données. Bref, il s'agit d'une analyse et d'une synthèse numérique et/ou graphique des données observées. Par exemple, la moyenne, le mode, la médiane, variance, la déviation standard (écart-type échantillonnal), le coefficient de Yule et Kendall et le coefficient d'asymétrie sont des indicateurs descriptifs à consigner dans des tableaux.

L'histogramme et le diagramme en barres sont des exemples de graphiques utilisés en statistique descriptive. La statistique analytique ou mathématique, quant à elle, permet de tirer des conclusions en se basant sur le calcul des probabilités. En d'autres termes, elle permet de tirer des conclusions par rapport à la population en se servant des mesures statistiques réalisées sur base des échantillons : c'est l'inférence statistique. Par exemple, la moyenne (μ) d'une variable ou la proportion (π) d'une population est estimée à partir de la moyenne $(\overline{x})$ ou de la proportion (p) observées dans l'échantillon. C'est dans cette partie de la statistique analytique que la théorie d'estimation et les tests d'hypothèse sont abordés. Pour faire de la statistique, il faut alors les données et ces dernières proviennent de l'enquête. Dans ce sens, la statistique part de l'enquête.

La statistique n'est pas seulement une science des données. Elle est une branche des mathématiques appliquées qui fait intervenir la théorie des probabilités (statistique mathématique) et l'informatique (data science). Pierre Dagnelie (Belgique, 1982), quant à lui, estime que *la statistique est composée de trois parties : la statistique administrative ou gouvernementale faite dans les institutions de statistique avec de grands ensembles de données, la statistique mathématique ou universitaire faite avec peu de données et qui a pour but la novation et la statistique appliquée ou de terrain faite dans les instituts de sondage d'opinion ou les facultés de médecine pour des problèmes concrets* [1]. La statistique, au même titre que les probabilités, est une science de l'aléatoire.

Le tableau 1 ci-après montre la différence entre la statistique et les probabilités.

Tableau 1 : Comparaison entre la statistique et les probabilités

Statistique	Probabilités
⇨ Elle part de l'expérience	⇨ Elles permettent de prédire ce qui va arriver après
⇨ Comportement réalisé	⇨ Comportement futur
⇨ Concepts pratiques	⇨ Concepts théoriques
⇨ Distribution des fréquences	⇨ Loi de probabilité
⇨ Polygone des fréquences cumulées	⇨ Fonction de répartition
⇨ Fréquence relative	⇨ Probabilité d'un événement
⇨ Moyenne arithmétique	⇨ Espérance mathématique ou moyenne espérée
⇨ Variable statistique	⇨ Variable aléatoire
⇨ Observations	⇨ Réalisations
⇨ Mécanisme déterministe	⇨ Mécanisme non déterministe (aléatoire ou

Source : Pr BARANKANIRA Emmanuel

Probabilités

Les concepts et les premières approches de la théorie des probabilités se définissent à partir des jeux de hasard [2]. La théorie des probabilités date du milieu du 17ème siècle avec Blaise Pascal (1623-1662) et Pierre Fermat (1601-1665). C'est au 19ème science que la théorie des probabilités s'est étendue au-delà des jeux de hasard avec entre autres auteurs Jacob Bernoulli (1654-1705) du Pays-Bas, Abraham de Moivre (1667-1754) d'Angleterre, Pierre Simon de Laplace (1749-1827) de France, Carl Gauss (1777-1805) d'Allemagne, Simon Poisson (1821-1894) de la France et Andreï Kolmogorov (1903-1987) de la Russie. La statistique et les probabilités sont indissociables. En statistique, par exemple, l'interprétation des intervalles de confiance et des résultats des tests statistiques requiert les notions de probabilité [3]. En probabilité, le calcul des moments à savoir l'espérance mathématique et la variance utilise des sommations utilisées en statistique.

Chapitre 1. Analyse combinatoire

Le comptage du nombre de façons de constituer des groupes *p* d'objets dans une liste de *n* objets requiert la connaissance des techniques permettant de faire le dénombrement. Ainsi par exemple, l'élection d'un délégué et de son adjoint (*p*=2) dans une classe de 50 étudiants (*n*=50) se fait en faisant recours à des arrangements du fait que la disposition est ordonnée. En effet, le délégué est élu en premier lieu. Ainsi, dans le cas où la notion d'ordre est importante, l'arrangement est utilisé, sinon il s'agit de la combinaison. Dans chacun des cas, il est possible de tenir compte du fait qu'il y a ou non des répétitions. Un cas particulier de l'arrangement est la permutation (n=p). En tant que branche des mathématiques, l'analyse combinatoire a pour objet de faire des dénombrements. De ce fait, elle étudie la manière de compter les éléments. Elle est fondée sur les notions d'arrangement et de combinaison. Elle est utilisée dans plusieurs branches des mathématiques impliquant la théorie des probabilités et les statistiques.

L'analyse combinatoire permet de résoudre les problèmes de la vie courante comme par exemples :

- De combien de façons peut-on asseoir 5 couples sur un banc de 10 places lors d'une fête de mariage si chaque couple doit rester ensemble ?
- Combien d'anagrammes peut-on écrire avec les lettres du mot « statistique » ?
- De combien de façons un port peut-il recevoir 3 bateaux si il ne dispose que de 4 places ?

1.1. Disposition ordonnée et non ordonnée

L'absence ou la présence de l'ordre dans le dénombrement est une notion importante. Cette souligne le fait que la position occupée par les éléments est prise en compte ou pas. Mathématiquement, la disposition est dite ordonnée si $(a,b) \neq (b,a)$ et non ordonnée si $(a,b) \equiv (b,a)$. Par exemple, lors de l'élection d'un délégué et de son adjoint dans une classe de 50 étudiants, il est clair que le délégué est élu en premier lieu et que le sous-délégué sera élu en deuxième lieu. Il en est de même lorsqu'un étudiant écrit son nom de famille : il commence par remplir la première case avant la deuxième. Il y a donc la notion d'ordre. Par contre, la constitution des groupes formés chacun de *p=3* étudiants choisis dans une liste de *n=5* étudiants ne fait pas intervenir la notion d'ordre : c'est la combinaison. L'arrangement, la permutation et la combinaison peuvent se faire avec ou sans répétition. La notion de répétition montre qu'un même élément peut être utilisé une ou plusieurs.

1.2. Principe de décomposition

Ce principe consiste à calculer le nombre de façons d'obtenir un résultat d'analyse combinatoire en faisant la décomposer d'une expérience en k épreuves élémentaires $n_1, \ldots, n_k$. Par exemple, pour calculer le nombre de plaques de voiture à fabriquer et donc le nombre de voitures qui seront servies si chaque plaque porte deux lettres différentes et doit être de la forme A123B sachant que le premier chiffre est différent de zéro, alors il y aura :

- 26 façons de choisir la première lettre ;
- 9 façons de choisir le premier chiffre ;
- 10 façons de choisir le deuxième chiffre ;
- 10 façons de choisir le troisième chiffre ;
- 25 façons de choisir la dernière lettre.

Selon ce principe, le nombre de plaques à fabriquer vaut 26×9×10×10×25=585000.

1.3. Arrangement

1.3.1. Arrangement sans répétitions

Un arrangement de p éléments distincts choisis parmi n éléments ($n \geq p$) est un groupement ou classement dans un ordre bien déterminé de p éléments choisis parmi les n éléments, chaque élément ne pouvant figurer qu'une seule fois dans chaque groupement :

$$A_n^p = \frac{n!}{(n-p)!} = \frac{n(n-1)(n-2)...(n-p+1)(n-p)!}{(n-p)!} = n(n-1)(n-2)...(n-p+1) \qquad (1.1)$$

Dans cette formule, il y a p facteurs.

Dans le cas où trois objets a, b et c sont donnés, les arrangements sans répétitions qu'il est possible de faire avec deux éléments choisis parmi les objets a, b et c sont : (a,b), (a,c), (b,a), (b,c), (c,a), (c,b).

Le nombre d'arrangements sans répétitions est donc :

$$A_3^2 = \frac{3!}{1!} = 6$$

Par exemple, dans une course de 6 chevaux, combien y a-t-il de podiums différents ?

La réponse à cette question nécessite la connaissance du nombre de places pour un podium. Un podium possède 3 places. Il est possible d'utiliser le principe de décomposition pour trouver la bonne réponse. En effet, il y a 6 possibilités de choisir la première place. Une fois cette place est choisie, il y aura alors 5 façons de choisir la deuxième place. Une fois que les deux premières places sont choisies, il y aura 4 possibilités de choisir la troisième place. Globalement, il y aura 6×5×4=120 possibilités.

Cette réponse peut aussi être trouvée par la formule de l'analyse combinatoire :

$$A_6^3 = \frac{6!}{3!} = \frac{6\times5\times4\times3!}{3!} = 6\times5\times4 = 120$$

Ce résultat est trouvé en faisant un produit de 3 (la valeur de p) facteurs.

Exercice :

Après les prolongations d'un match de football, y a-t-il combien de manières de choisir un tireur de pénalités ?

Sachant que le match de football comporte onze joueurs et que tirs sont faits, la réponse à cette question est que le nombre de manières de choisir un tireur de pénalités parmi les onze joueurs est :

$$A_{11}^5 = \frac{11!}{6!} = \frac{11\times10\times9\times8\times7\times6!}{6!} = 11\times10\times9\times8\times7 = 55440$$

1.3.2. Arrangement avec répétitions

Un arrangement (avec répétitions) de p éléments choisis parmi n éléments ($n \geq p$) est un groupement ou classement dans un ordre bien déterminé de p éléments choisis parmi les n éléments, chaque élément pouvant figurer plusieurs fois dans chaque groupement :

$$\overline{A}_n^p = n^p \tag{1.2}$$

Les arrangements avec répétitions de deux objets choisis deux à deux parmi les trois objets a, b et c choisis sont: (a,a), (a,b), (a,c), (b,b), (b,a), (b,c), (c,c), (c,a), (c,b).

Le nombre d'arrangements avec répétitions est donc :

$$A_3^2 = 3^2 = 9$$

Exercice :

Combien de plaques de voitures différentes est-il possible de faire sachant que chaque plaque porte 3 lettres et que les lettres A, B, C, D, E, F, G sont à votre disposition ?

La réponse à cette question est que le nombre de plaques qu'il est possible de fabriquer est donné par la formule :

$$A_7^3 = 7^3 = 343$$

Cette réponse peut aussi être trouvée en utilisant le principe de la décomposition ci-haut décrite.

1.4. Permutation

1.4.1. Permutation sans répétitions

Dans le cas où le n est égal à p dans la formule de l'arrangement sans répétitions, alors il s'agit d'une permutation. Autrement dit, une permutation est une disposition ordonnée dans laquelle le nombre de places à pourvoir est égal au nombre d'objets à grouper.

Une permutation de n objets est un arrangement de n objets pris n à n :

$$P_n = A_n^n = n! = n \times (n-1) \times (n-2) \times (n-3) \times \ldots \times 3 \times 2 \times 1 \tag{1.3}$$

Par convention, $0! = 1$. De plus, de manière récursive, il vient que $n! = n(n-1)! = n(n-1)(n-2)! = n(n-1)(n-2)(n-2)! = \ldots = \prod_{i=1}^{n} i$.

1.4.2. Permutation avec répétitions

Considérons n objets parmi lesquels il y a n_1 objets de type 1, n_2 objets de type 2, n_3 objets de type 3, …, et n_k objets de type k. Dans ce cas, il y aura n permutations parmi lesquels n_1 permutations des objets de type 1, n_2 permutations des objets de type 2, n_3 permutations des objets de type 3, …, et n_k permutations des objets de type k.

Sachant que $n = \sum_{i=1}^{k} n_i$, le nombre de permutations de n objets avec répétitions s'écrit :

$$\overline{P}_n = \frac{n!}{n_1!\,n_2!\,n_3!\ldots n_k!} \tag{1.4}$$

Exemple :

Considérons le mot CAOUTCHOUC. La lettre C est répété trois fois, la lettre A une fois, la lettre O deux fois, la lettre U deux fois, la lettre T une fois et la lettre H une fois. Le nombre de mots qu'il est possible de fabriquer avec les lettres de ce mot vaut :

$$\overline{P}_{10} = \frac{10!}{3!1!2!2!1!1!} = \frac{3628800}{24} = 151200$$

1.5. Combinaison

1.5.1. Combinaison sans répétitions

La combinaison est utilisée dans le cas où la notion d'ordre n'a pas d'importance. La combinaison sans répétitions de n objets pris p à p est le nombre de combinaisons de ces n objets qu'il est possible de faire sachant que chaque combinaison est formée de p objets et que chaque objet figure une seule fois dans chaque combinaison. Ce nombre de combinaisons est donné par :

$$C_n^p = \frac{n!}{p!(n-p)!} = \frac{A_n^p}{p!} \tag{1.5}$$

Dans le cas de trois objets a, b et c, les combinaisons sans répétitions de trois objets choisis deux à deux sont: (a,b), (a,c), (b,c) et leur nombre vaut :

$$C_3^2 = \frac{3!}{2!1!} = 3$$

1.5.2. Combinaison avec répétitions

Dans le cas où il y a des répétitions, la combinaison de n objets pris p à p est le nombre de combinaisons de ces n objets qu'il est possible de faire sachant que chaque combinaison est formée de p objets et que chaque objet figure plusieurs fois dans chaque combinaison. Ce nombre de combinaisons est donné par :

$$K_n^p = C_{n+p-1}^p = \frac{(n+p-1)!}{p!(n-1)!} \tag{1.6}$$

Dans le cas de trois objets a, b et c, les combinaisons avec répétitions de trois objets choisis deux à deux sont: (a,a), (a,b), (a,c), (b,b), (b,c), (c,c), et leur nombre vaut :

$$\overline{K}_3^2 = C_4^2 = \frac{4!}{2!2!} = 6$$

La combinaison vérifie entre autres propriétés :

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$
$$C_n^p = C_n^{n-p}$$
$$C_n^0 = C_n^n = 1$$

$$C_{k+l}^n = \sum_{i=0}^{n} C_k^i C_l^{n-i}, 0 \le n \le min(k,l)$$

$$C_{k+l}^n = \sum_{i=max(0,n-l)}^{min(n,k)} C_k^i C_l^{n-i}, n \le k+l \ge 1 \quad ()$$

et le binôme de Newton.

1.5.3. Binôme de Newton

Le binôme de Newton est une application de la formule de la combinaison sans répétitions. En effet, les coefficients des développements $(a+b)^n$ sont obtenus à l'aide de la combinaison et sont appelés des coefficients binomiaux.

$$(a+b)^1 = a+b = C_1^0 a + C_1^1 b$$

$$(a+b)^2 = a^2 + 2ab + b^2 = C_2^0 a^2 + C_2^1 ab + C_2^2 b^2$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 = C_3^0 a^3 + C_3^1 a^2 b + C_3^2 ab^2 + C_3^3 b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 = C_4^0 a^4 + C_4^1 a^3 b + C_4^2 a^2 b^2 + C_4^3 ab^3 + C_4^4 b^4$$

…

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k} \quad (1.7)$$

Prouvons cela par récurrence :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b)\sum_{k=0}^{n} C_n^k a^k b^{n-k} \\
&= \sum_{k=0}^{n} C_n^k a^{k+1} b^{n-k} + \sum_{k=0}^{n} C_n^k a^k b^{n+1-k} \\
&= \sum_{j=1}^{n+1} C_n^{j-1} a^j b^{n+1-j} + \sum_{k=0}^{n} C_n^k a^k b^{n+1-k} \textit{ en posant } j=k+1 \\
&= \left(\sum_{k=1}^{n} C_n^{k-1} a^k b^{n+1-k} + C_n^n a^{n+1} b^0\right) + \left(C_n^0 a^0 b^{n+1} + \sum_{k=1}^{n} C_n^k a^k b^{n+1-k}\right) \\
&= \left(\sum_{k=1}^{n} C_n^{k-1} a^k b^{n+1-k} + a^{n+1}\right) + \left(b^{n+1} + \sum_{k=1}^{n} C_n^k a^k b^{n+1-k}\right) \\
&= a^{n+1} + \sum_{k=1}^{n} \left(C_n^{k-1} + C_n^k\right) a^k b^{n+1-k} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} C_{n+1}^k a^k b^{n+1-k} + b^{n+1} \\
&= \sum_{k=0}^{n+1} C_{n+1}^k a^k b^{n+1-k}
\end{aligned}$$

En utilisant la propriété de symétrie, cette formule du binôme de Newton devient :

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k \tag{1.8}$$

Les coefficients binomiaux, obtenus à l'aide de la programmation sous le logiciel R pour n=9, sont consignés dans le triangle de Pascal ci-après [4] :

```
                  1
                1   1
              1   2   1
            1   3   3   1
          1   4   6   4   1
        1   5  10  10   5   1
      1   6  15  20  15   6   1
    1   7  21  35  35  21   7   1
  1   8  28  56  70  56  28   8   1
1   9  36  84 126 126  84  36   9   1
```

Figure 1 : Triangle de Pascal
Source : Pr BARANKANIRA Emmanuel

1.6. Exercices d'application

1° Montrez que :

a) $C_n^k = C_n^{n-k}$

b) $C_n^r = C_{n-1}^r + C_{n-1}^{r-1}$

c) $C_n^0 + C_n^1 + C_n^2 + \cdots + C_n^n = 2^n$

d) $C_r^{n+1} = C_{r-1}^n + C_r^n$

e) $C_{n-1}^{n+r-1} = C_r^{n+r-1}$

2° Calculez ou simplifiez :

a) $\dfrac{n!}{(n-1)!}$ b) $\dfrac{(n+2)!}{n!}$ c) $\dfrac{(n-1)!}{(n+2)!}$ d) $\dfrac{(n-r+1)!}{(n-r-1)!}$ e) $5!$ f) $\dfrac{1}{n!} + \dfrac{1}{(n-1)!}$ g) $\dfrac{13!}{11!}$

3° Trouvez les développements de : a) $(a-2r)^7$ b) $(x-2y)^5$

4° Simplifiez et calculez les nombres suivants :

a) $\dfrac{8!}{11!}$ b) $\dfrac{5!-8!}{4!-7!}$ c) $\dfrac{5!6!}{4!7!}$

5° Trouvez le 5ème terme de $(x^3+2)^{17}$

Chapitre 2. Calcul des probabilités

2.1. Expérience aléatoire

Avant de définir la probabilité, il convient de parler de ce qu'est un phénomène aléatoire. Un phénomène est dit aléatoire lorsqu'il ne se déroule pas de la même manière une fois qu'il est répété plusieurs fois. Autrement dit, le résultat de ce phénomène change de façon aléatoire (le hasard intervient) et est imprévisible. Ce phénomène peut porter le nom d'expérience, c'est-à-dire un processus d'obtention d'un résultat (outcome) [5].

Gilbert SAPORTA définit une expérience aléatoire comme suit :

« Une expérience est qualifiée d'aléatoire ou de stochastique si nous ne pouvons pas prévoir son résultat et, si répétée dans des conditions identiques, elle peut donner lieu à des résultats différents. » [6].

Une expérience aléatoire (stochastique), appelée également épreuve ou tirage, est donc une expérience pour laquelle :

- il est possible de décrire l'ensemble de tous les résultats possibles ;
- il n'est pas possible de prédire avec certitude les résultats ;
- il est possible de répéter l'expérience à volonté et dans les mêmes conditions et obtenir les mêmes résultats.

Il s'agit par exemples du lancer d'une pièce de monnaie, du jeu de cartes, du jet de dé, du jeu de football et du tirage d'une personne. Si l'expérience n'est pas aléatoire, alors elle déterministe. Le tableau 2 montre quelques exemples de la vie courante concernant les expériences aléatoires et leurs résultats possibles.

Tableau 2 : Expérience aléatoire et résultats possibles

Expérience aléatoire	Résultats possibles
Jeter une pièce de monnaie	Pile, Face
Lancer deux pièces de monnaie	PP, PF, FP, FF
Choisir une carte	Rouge, Noire
Choisir une carte	Carreau, Cœur, Pique, Trèfle
Jouer au football	Gagné, Perdu, Égalité
Inspecter une pièce	Défectueuse, En bon état
Observer le genre	Homme, Femme
Examen de Probabilités et statistique analytique	Succès, Échec
Tirer une personne	Malade, Pas malade

Source : Pr BARANKANIRA Emmanuel

Ces résultats sont mutuellement exclusifs lorsqu'ils ne peuvent pas se réaliser en même temps (par exemple Homme et Femme lorsque c'est le genre qui est observé) et exhaustifs lorsqu'un des résultats possible se réalise (par exemple, lorsqu'une femme est enceinte, elle mettra au monde soit un garçon, soit une fille et pas les deux en un seul coup).

2.2. Espace échantillonnal

Lorsqu'une pièce de monnaie est lancée, la face PILE (qui indique la valeur de la monnaie) et la face FACE (qui indique une image) sont observées. Ainsi, les résultats possibles de l'expérience aléatoire qui consiste à jeter une pièce de monnaie sont PILE ou FACE. Ces résultats sont les éléments de l'espace échantillonnal, appelé également ensemble fondamental, espace d'échantillonnage, espace des réalisations, ensemble élémentaire ou univers des possibles et noté généralement par S ou par Ω. Dans la suite, la notation Ω sera privilégiée. Dans le cas du jet d'une pièce de monnaie, l'écriture en extension de l'univers $\Omega = \{P, F\}$. Cet ensemble est fini. L'univers peut être fini, infini, continu ou discontinu (discret). Par exemple, lors du jeu de cartes, la carte peut être soit rouge ou noire, soit un carreau, un cœur, un pique ou une trèfle, soit encore un as, une dame, une roi, etc. Lors du jet de dé parfait, (équilibré, non pipé, non truqué), il est possible d'observer les faces 1, 2, 3, 4, 5, ou 6 et après un match de football, le statut sera gagné, perdu ou égalité. Pour le tirage d'une personne, cette dernière peut être par exemple un homme ou une femme, d'où $\Omega = \{Homme, Femme\}$.

Le cardinal d'un ensemble est le nombre d'éléments que cet ensemble contient. Par exemple, lorsqu'une pièce de monnaie qui est supposée être non truquée (parfaite) est lancée, les faces PILE et FACE sont observées. Dans ce cas, $\Omega = \{P, F\}$ et $card(\Omega) = \#\Omega = 2$. Remarquons que le face PILE est la face qui indique la valeur de la pièce de monnaie. De même, lorsque cette pièce est lancée deux fois, ce qui est équivalent au fait que deux pièces de monnaie sont lancées en même temps, alors $\Omega = \{(P,P),(P,F),(F,P),(F,F)\}$ et $card(\Omega) = \#\Omega = 4$. Dans ces deux cas, l'ensemble des résultats possible de cette expérience aléatoire, appelées aussi issues ou éventualités, sont obtenues à l'aide d'une liste.

Une autre façon d'obtenir ces résultats est de construire une table de contingence ou un arbre de décision comme le montrent le tableau 3 et la figure 2.

Tableau 3 : Table de contingence

1ère pièce	2ème pièce		
	P	F	Total
P	PP	PF	PP, PF
F	FP	FF	FP, FF
Total	PP, FP	PF, FF	Ω

Source : Pr BARANKANIRA Emmanuel

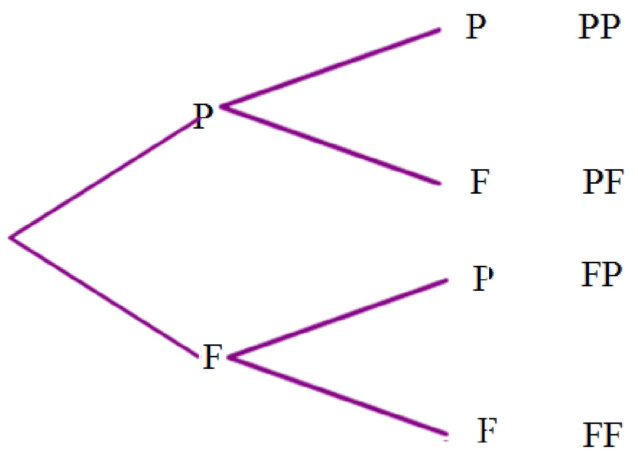


Figure 2 : Arbre de décision
Source : Pr BARANKANIRA Emmanuel

2.3. Événement aléatoire

Une expérience est tout ensemble de résultats possibles et une partie de cet ensemble est un événement. En d'autres mots, un événement est un résultat possible d'une épreuve. Il s'agit d'une éventualité susceptible d'être réalisée lors de l'expérience aléatoire. Sa réalisation dépendant du hasard. De ce fait, l'événement est dit aléatoire [7].

Un événement peut être simple ou élémentaire (résultat avec une seule caractéristique), composé (ensemble de résultats ou d'événements simples, avec deux caractéristiques ou plus) ou joint (cas particulier de 2 événements se réalisant en même temps).

2.3.1. Événement élémentaire

En théorie des probabilités, un événement élémentaire est constitué d'un seul élément. Par exemple, lorsqu'un dé (il a six faces) non pipé est lancé, l'ensemble fondamental est $\Omega = \{1,2,3,4,5,6\}$. Cela permet de définir six événements élémentaires qui correspondant, chacun, à l'apparition d'un numéro de la face, c'est-à-dire 1, 2, 3, 4, 5, 6. L'obtention d'un numéro de la face est un événement élémentaire.

2.3.2. Événement composé

Dans un jeu, il est possible de considérer l'obtention de plusieurs résultats. Dans ce cas, l'événement qui représente cette obtention est un événement composé. Par exemple, lorsqu'un dé est lancé, l'événement A : « Obtenir un nombre pair » est donné en extension par $A = \{2,4,6\}$. Cet événement est dit composé.

Il existe plusieurs autres façons de considérer un événement composé. Par exemples, quand deux pièces de monnaie sont lancées en même temps ou qu'une pièce de monnaie est lancée deux fois, les événements de cette expérience aléatoire peuvent être :

A : « Obtenir au moins une fois PILE » $\Longrightarrow$ A={PP, FP, PF}

B : « La première pièce amène PILE et l'autre FACE » $\Longrightarrow$ B={PF, FP}

C : « La première pièce amène FACE » $\Longrightarrow$ C={ FF, FP}

D : « Les deux pièces amènent FACE » $\Longrightarrow$ F={ FF}

Pour les événements composés, différentes opérations mathématiques telles que la réunion (appelée aussi somme), l'intersection (appelée aussi produit), la complémentation et la différence peuvent être effectuées afin d'obtenir des événements composés :

- $A \cup B$: se lit « A ou B », ce qui veut dire que l'un au moins des deux événements est réalisé ;
- $A \cap B$: se lit « A et B », ce qui veut dire que les deux événements se réalisent en même temps ;
- A^c : se lit « complémentaire de A », ce qui veut dire que l'événement A^c se réalise si A n'est pas réalisé ;
- $A - B = A \cap B^c$: se lit « A moins B », ce qui veut dire que A s'est réalisé, mais pas B.

Il existe d'autres événements tels que :

- ϕ : événement impossible, par exemple « Obtenir le chiffre 10 lors du lancé d'un dé non pipé ». Cet événement n'a aucune chance de se produire ;
- Ω : événement certain, par exemple, « Obtenir PILE ou FACE » lors d'un jet d'une pièce de monnaie ;
- $A \cap B = \phi$: événements incompatibles ou mutuellement exclusifs. C'est le cas, par exemple, des événements A : « Obtenir un nombre pair » et B : « Obtenir un nombre pair ». En termes ensemblistes, cela s'écrit : $A \cap B = \{2,4,6\} \cap \{1,3,5\} = \phi$.

2.4. Notion de probabilité

2.4.1. Cardinal d'un ensemble

Il existe un lien entre les notations ensemblistes et les notations probabilistes. Ces notations permettent de calculer le nombre d'éléments formant un ensemble (son cardinal). À titre exemplatif, considérons un ensemble formé de deux éléments : $E = \{a,b\}$. Dans ce cas, son cardinal vaut 2 et se note : $card(E) = \# E = 2$. En partant de cet ensemble, l'ensemble des parties de cet ensemble est formé de l'ensemble vide, de tous les singletons, de toutes les paires et de l'ensemble plein :

$$P_E = P(E) = \{\phi, \{a\}, \{b\}, \{a,b\}\}$$

Ainsi, son cardinal vaut :

$$Card\left(P_E\right)=2^2=2^{card(E)}=4\,.$$

De même, considérons un ensemble formé de trois éléments : $E=\{a,b,c\}$.

L'ensemble des parties de cet ensemble est :

$$P_E=\{\phi,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},\{a,b,c\}\}$$

et son cardinal vaut :

$$Card\left(P_E\right)=2^3=2^{card(E)}=8$$

Le couple $\left(\Omega,P_\Omega\right)$ est appelé espace probabilisable, c'est-à-dire un espace qui peut être muni d'une probabilité. P_Ω désigne l'ensemble de tous les événements [8,9]. Dans la plupart des cas, P_Ω est remplacé par une partie non vide de Ω qui est considérée comme étant stable par passage au complémentaire et par union dénombrable, ce qui constitue une tribu ou une $\sigma-alg\grave{e}bre$ A.

Le cardinal de la réunion de trois ensembles s'écrit :

$$\begin{aligned}
card\left(A\cup\overbrace{B\cup C}^{D}\right)&=Card\left(A\cup D\right)\\
&=card\left(A\right)+card\left(D\right)-card\left(A\cap D\right)\\
&=card\left(A\right)+card\left(B\cup C\right)-card\left[A\cap\left(B\cup C\right)\right]\\
&=card\left(A\right)+card\left(B\right)+card\left(C\right)-card\left(B\cap C\right)-card\left[\left(A\cap B\right)\cup\left(A\cup C\right)\right]\\
&=card\left(A\right)+card\left(B\right)+card\left(C\right)-card\left(B\cap C\right)-card\left(A\cap B\right)-card\left(A\cap C\right)+card\left(A\cap B\cap C\right)\\
&=card\left(A\right)+card\left(B\right)+card\left(C\right)-card\left(A\cap B\right)-card\left(A\cap C\right)-card\left(B\cap C\right)+card\left(A\cap B\cap C\right)
\end{aligned}$$

Considérons un diagramme de **Venn** donné par la figure 3. Le diagramme de Venn est une autre façon de montrer l'ensemble des résultats possibles d'une expérience aléatoire.

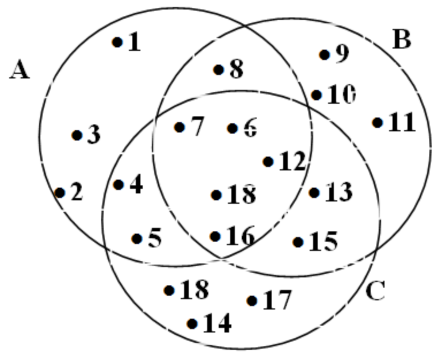


Figure 3 : Cardinal d'un ensemble
Source : Pr BARANKANIRA Emmanuel

Les cardinaux des ensembles *A*, *B* et *C* sont respectivement :

$$card(A)=11 \quad card(B)=11 \quad card(C)=12$$

Les cardinaux des ensembles intersection sont respectivement :

$$card(A\cap B)=6 \quad card(A\cap C)=7 \quad card(B\cap C)=7 \quad card(A\cap B\cap C)=5$$

Les cardinaux des ensembles réunion sont respectivement :

$$card(A\cup B)=15 \quad card(A\cup C)=15 \quad card(B\cup C)=15 \quad card(A\cup B\cup C)=18$$

Le cardinal de la réunion des ensembles *A* et *B* est :

$$card(A\cup B)=card(A)+card(B)-card(A\cap B)=10+11-6=15$$

Le cardinal de la réunion des trois ensembles *A, B* et *C* est :

$$\begin{aligned} card(A\cup B\cup C) &= \#A+\#B+\#C-\#(A\cap B)-\#(A\cap C)-\#(B\cap C)+\#(A\cap B\cap C) \\ &= 10+11+12-6-7-7+5 \\ &= 38-20 \\ &= 18 \end{aligned}$$

2.4.2. Notations ensemblistes et probabilistes

Le tableau 4 ci-après montre la correspondance entre les notations ensemblistes et les notations probabilistes.

Tableau 4 : Notations ensemblistes et notations probabilistes

Notations ensemblistes	Notations probabilistes
ϕ : Ensemble vide	ϕ : Événement impossible
$A\cap B$	A et B sont réalisés simultanément
$A\cap B\neq\phi$: Ensembles non disjoints	$A\cap B\neq\phi$: Événements compatibles
$A\cap B=\phi$: Ensembles disjoints	$A\cap B=\phi$: Événements incompatibles
$A\subset B$: A est inclus dans B	$A\subset B$: L'événement A implique B
Ω	Événement certain
$A\cup B$	Au moins un des 2 événements est réalisé
$\overline{A}=A^c=\Omega\backslash A$	L'événement A n'est pas réalisé

Source : Jean-Pierre Lecoutre, 2016

2.4.3. Définition de la probabilité

La notion de probabilité peut se définir de trois manières différentes. Premièrement, la probabilité se définit comme étant le rapport entre le nombre de cas favorables à un événement et le nombre de cas possibles, c'est-à-dire le cardinal de l'univers [7] :

$$p(A) = \frac{nombre\ de\ cas\ farobales\ à\ A}{nombre\ de\ cas\ possibles} = \frac{\# A}{\# \Omega} \tag{2.1}$$

Cette définition est appelée définition de Laplace. Longtemps considérée comme définition classique, « *la probabilité d'un événement est définie comme le nombre de cas favorables pour un événement donné divisé par le nombre total d'issues possibles à l'expérience aléatoire* » [10].

Deuxièmement, la probabilité d'un événement se calcule en utilisant la notion de limite. En effet, lorsqu'une expérience aléatoire est répétée n fois et que l'événement d'intérêt se réalise f fois, alors la probabilité de réalisation de l'événement A se calcule comme suit :

$$p(A) = \lim_{n \to +\infty} \frac{f}{n} \tag{2.2}$$

Troisièmement, la probabilité est définie comme étant une application dont l'ensemble de départ est l'univers et l'ensemble d'arrivée est $[0, 1]$:

$$\begin{aligned} p : \Omega &\to [0, 1] \\ A &\mapsto p(A) \end{aligned}$$

telle que les axiomes suivants, qui constituent l'axiomatique de Kolmogorov, soient vérifiés :

a) $0 \leq p(A) \leq 1$

b) $p(\emptyset) = 0$, $p(\Omega) = 1$

c) $p(A \cup B) = p(A) + p(B) - p(A \cap B)$ si A et B sont compatibles

$p(A \cup B) = p(A) + p(B)$ si A et B sont incompatibles

Cette dernière relation vient du fait que deux événements A et B sont incompatibles (ou mutuellement exclusifs) ssi :

$$p(A \cap B) = 0 \tag{2.3}$$

Dans le cas de trois événements, la probabilité jointe de trois événements non exclusifs s'écrit [11]:

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

La généralisation de la probabilité de réalisation simultanée de k événements qui sont mutuellement non exclusifs est :

$$p\left(\bigcup_{i=1}^{k} A_i\right)=\sum_{i=1}^{k} p(A_i)-p\left(\bigcap_{i=1,j=1;j\neq i}^{k}\left(A_i\cap A_j\right)\right)+p\left(\bigcap_{i=1}^{k} A_i\right) \tag{2.4}$$

Dans le cas d'indépendance, cette formule devient :

$$P(A_1\cup A_2\cdots\cup A_k)=\sum_{i=1}^{k} P(A_i) \tag{2.5}$$

C'est l'**égalité de Poincaré.**

De plus, en l'absence d'indépendance, il vient l'**inégalité de Boole** :

$$p(A_1\cup A_2\cdots\cup A_k)\leq\sum_{i=1}^{k} p(A_i) \tag{2.6}$$

Sachant que $\Omega=A\cup\bar{A}$, $p\left(A\cup\bar{A}\right)=P(\Omega)=1$ et $p\left(A\cap\bar{A}\right)=p(\phi)=0$, alors le développement donne :

$$P\left(\bar{A}\right)=1-P(A) \tag{2.7}$$

Il est à remarquer que les événements dont la probabilité de réalisation est identique sont dits **équiprobables**. C'est par exemples le cas de l'obtention d'un numéro de face lorsqu'un dé non truqué ou une pièce de monnaie sont lancés. Cela n'est pas le même cas lorsque trois boules sont tirées d'une urne contenant **8 boules rouges**, **3 blanches** et **9 bleues**.

En effet, considérons les événements suivants :

A : « Obtenir trois boules rouges »
B : « Obtenir trois boules blanches »
C : « Obtenir trois boules bleues »
Les probabilités de ces événements sont

$$P(A)=\frac{C_8^3}{C_{20}^3}=\frac{\frac{8!}{3!5!}}{\frac{20!}{3!17!}}=\frac{8!}{5!}\frac{17!}{20!}=\frac{8\times7\times6}{20\times19\times18}=\frac{2\times7}{5\times19\times3}\simeq0,05.$$

$$P(B)=\frac{C_3^3}{C_{20}^3}=\frac{\frac{3!}{3!0!}}{\frac{20!}{3!17!}}=\frac{3!}{\frac{20!}{17!}}=\frac{3!17!}{20!}=\frac{6}{20\times19\times18}=\frac{1}{20\times19\times3}\simeq0,00$$

$$p(C)=\frac{C_9^3}{C_{20}^3}=\frac{\frac{9!}{3!6!}}{\frac{20!}{3!17!}}=\frac{9!}{6!}\frac{17!}{20!}=\frac{9\times8\times7}{20\times19\times18}=\frac{9\times7}{5\times19\times9}\simeq 0{,}07$$

Le couple (Ω ,A) est appelé **espace probabilisable**, c'est-à-dire un espace qui peut être muni d'une probabilité. A est une partie de l'ensemble des parties de l'univers telle que la réunion et l'intersection soient stables dans A. Le triplet (Ω ,A, p) est appelé **espace probabilisé**, c'est-à-dire un espace qui est muni d'une probabilité. Cet espace décrit de façon complète une épreuve. La fonction p est une mesure de probabilités définie sur un espace des événements.

2.4.4. Incompatibilité en probabilité

Deux événements A et B sont dits incompatibles en probabilité lorsque la probabilité de réalisation simultanée est nulle. Cela veut dire que, mathématiquement, leur intersection est vide ou qu'ils ne peuvent pas se réaliser en même temps :

$$p(A\cap B)=0 \tag{2.8}$$

Autrement dit, si $A\cap B=\phi$, c'est-à-dire si la réalisation simultanée des événements A et B est impossible, alors les événements A et B sont incompatibles [6]. Ces deux événements sont dits exclusifs. L'exclusivité ou l'incompatibilité mutuelle entre trois événements s'écrit $p(A\cap B\cap C)=0$.

Dans le cas d'incompatibilité, la relation $p(A\cup B)=p(A)+p(B)-p(A\cap B)$ devient :

$$p(A\cup B)=p(A)+p(B) \tag{2.9}$$

2.4.5. Indépendance en probabilité

Deux événements A et B sont dits indépendants en probabilité lorsque la probabilité de réalisation simultanée est égale au produit les probabilités individuelles. Mathématiquement, cela veut dire que :

$$p(A\cap B)=p(A)\,p(B) \tag{2.10}$$

Dans le cas de plusieurs événements, l'indépendance mutuelle et l'indépendance deux à deux donnent :

$$p(A_1\cap A_2\cap A_3\cap\cdots\cap A_k)=p(A_1)\,p(A_2)\,p(A_3)\ldots p(A_k)=\prod_{i=1}^{k}p(A_i) \tag{2.11.a}$$

ou de façon condensée :

$$p\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} p(A_i) \qquad (2.11.b)$$

2.5. Probabilité conditionnelle

L'une des notions importantes en théorie des probabilités est la notion de probabilité conditionnelle. En effet, le calcul des probabilités de réalisation des événements donnés lorsqu'un autre événement s'est déjà réalisé se fait par la connaissance de l'information donnée par ce dernier événement d'une part. Dans ce cas, les probabilités cherchées sont des probabilités conditionnelles. D'autre part, même lorsqu'aucune information partielle (a priori) n'est disponible, il est souvent utile d'utiliser un détour pour certaines probabilités conditionnelles afin de faciliter le calcul des probabilités cherchées.

De façon plus précise, considérons deux événements *A* et *B*, en supposant que la probabilité que l'événement *B* se réalise ne soit pas nulle. Dans ce cas, la probabilité de réalisation de l'événement *A* sachant que l'événement *B* s'est réalisé est, par définition, une probabilité conditionnelle donnée par :

$$p(A/B) = \frac{p(A \cap B)}{p(B)} \qquad (2.12)$$

A/B se lit *A* sachant *B* ou *A* si *B* et peut aussi se noter $A|B$. Dans la suite, nous allons adopter la notation *A/B*.

La relation (2.12) équivaut à dire que :

$$p(A \cap B) = p(A/B)\,p(B) \qquad (2.13)$$

De même :

$$p(B/A) = \frac{p(B \cap A)}{p(A)} \Rightarrow p(B \cap A) = p(B/A)\,p(A) \qquad (2.14)$$

Grâce à la commutativité de l'intersection et vu les relations (2.13) et (2.14), le théorème de l'inversion des probabilités stipule que :

$$p(B/A) = P(A/B)\,P(B) \qquad (2.15)$$

Puisque $A = (A \cap B) \cup (A \cap \overline{B})$ et étant donné l'incompatibilité des événements, le théorème des probabilités totales s'écrit :

$$p(A) = p\left[(A \cap B) \cup (A \cap \overline{B})\right] = p(A \cap B) + p(A \cap \overline{B})$$

ou encore, en utilisant la règle de multiplication:

$$p(A) = p(A / B)\, p(B) + p(A / \overline{B})\, p(\overline{B})$$

Les relations (2.11), (2.12) et (2.13) permettent d'écrire **la règle de multiplication** suivante :

$$p(A \cap B) = p(A / B)\, p(B) = p(B / A)\, p(A) \quad (2.16)$$

La généralisation de cette règle conduit à :

$$p(A_1 \cap A_2 \cap \cdots \cap A_k) = p(A_1)\, p(A_2/A_1)\, p\left[A_3/(A_1 \cap A_2)\right] \cdots p\left[A_k / (A_1 \cap A_2 \cap \ldots \cap A_{k-1})\right]$$

Comme conséquence de l'indépendance entre les événements, il vient :

$$p(A/B) = \frac{p(A \cap B)}{P(B)} = \frac{p(A)\, p(B)}{p(B)} = P(A) \quad (2.17.a)$$

$$p(B/A) = \frac{p(B \cap A)}{p(A)} = \frac{p(B)\, p(A)}{p(A)} = p(B) \quad (2.17.b)$$

Cela veut dire que la probabilité de réalisation d'un événement sachant qu'un autre événement qui lui est indépendant s'est réalisé ne change pas la probabilité de réalisation de cet événement.

La probabilité conditionnelle vérifie entre autres propriétés :

$$p(\overline{A}/B) = 1 - p(A/B) \quad (2.18)$$

$$p(\Omega / B) = p(\Omega) = 1 \quad (2.19)$$

Considérons $(A_i)_{1 \le i \le n}$ une famille d'événements mutuellement exclusifs. La probabilité de la réunion de ces événements sachant un événement donné B vaut :

$$p\left[\left(\bigcup_{i=1}^{k} A_i\right) / B\right] = \frac{p\left[\left(\bigcup_{i=1}^{k} A_i\right) \cap B\right]}{p(B)} = \frac{p\left[\bigcup_{i=1}^{k} (A_i / B)\right]}{p(B)} = \frac{\sum_{i=1}^{k} p(A_i / B)}{p(B)} \quad (2.20)$$

La probabilité conditionnelle peut se calculer soit en utilisant la définition, soit en considérant la probabilité du premier événement par rapport à l'espace échantillonnal.

Exemple : Jeux de cartes

Deux cartes sont tirées sans remise successivement dans un jeu de 52 cartes. La question est de déterminer la probabilité de tirer un as au deuxième tirage sachant qu'il a été obtenu au premier tirage [7]. Il est à noter que dans ce jeu, il y a :

- 26 cartes noires et 26 cartes rouges
- 13 carreaux, 13 cœurs, 13 piques, 13 trèfles
- Pour chacun de ces 4 groupes, il y a 1 as, 1 roi, 1 dame, 1 valet, 10, 9, 8, 7, 6, 5, 4, 3, 2
- 4 as et 48 non as

Pour résoudre le problème, considérons les événements suivants :

A_1 : « Obtenir un As au premier tirage »

A_2 : « Obtenir un As au deuxième tirage »

Les probabilités de ces événements sont :

$$p(A_1) = \frac{4}{52}$$
$$p(A_2) = \frac{3}{51}$$

La formule de la probabilité conditionnelle donne :

$$p(A_2 / A_1) = \frac{p(A_2 \cap A_1)}{p(A_1)} = \frac{p(A_1)\,p(A_2 / A_1)}{p(A_1)} == \frac{\frac{4}{52} \times \frac{3}{52}}{\frac{4}{52}} = \frac{3}{51} = \frac{1}{17}$$

Exercice :

Dans une urne qui contient 2 boules rouges et 3 boucles noires, 4 personnes T, U, V et W tirent cet ordre successivement une boule de l'urne sans la remettre. La 1ère personne qui tire une boule rouge gagne le jeu. Calculer la probabilité de gain de chaque personne.

Solution

Soient les événements :

T_1 : « La personne T tire une boule rouge au 1er tirage »

U_1 : « La personne U tire une boule rouge au 1er tirage »

V_1 : « La personne V tire une boule rouge au 1er tirage »

W_1 : « La personne W tire une boule rouge au 1er tirage »

T_2 : « La personne T tire une boule noire au 2ème tirage »

U_2 : « La personne U tire une boule noire au 2ème tirage »

V_2 : « La personne V tire une boule noire au $2^{ème}$ tirage »

W_2 : « La personne W tire une boule noire au $2^{ème}$ tirage »

T : « La personne T gagne le jeu »

U : « La personne U gagne le jeu »

V : « La personne V gagne le jeu »

W : « La personne W gagne le jeu »

La probabilité de gain de la personne T vaut :

$$p(T) = p(T_1) = \frac{2}{5} = 0,40$$

La probabilité de gain de la personne U vaut :

$$p(U) = p(T_2 \cap U_1) = p(T_2)\, p(U_1 / T_2) = \frac{3}{5} \times \frac{2}{4} = \frac{6}{20} = 0,30$$

La probabilité de la personne V vaut :

$$p(V) = p(T_2 \cap U_2 \cap V_1) = p(T_2)\, p(U_2 / T_2)\, p\left[V_1 / (T_2 \cap V_2)\right] = \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{1}{5} = 0,20$$

La probabilité de gain de la personne W vaut :

$$p(W) = p(T_2 \cap U_2 \cap V_2 \cap W_1) = p(T_2)\, p(U_2 / T_2)\, p\left[V_2 / (T_2 \cap U_2)\right] p\left[W_1 / (T_2 \cap U_2 \cap V_2)\right]$$
$$= \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \times \frac{2}{2} = \frac{1}{10} = 0,10$$

2.6. Théorème de Bayes

Considérons A_1, A_2, A_k k événements mutuellement exclusifs constituant l'univers des possibles. Dans ce cas, cet univers s'écrit [12] :

$$\Omega = \bigcup_{i=1}^{k} A_i \tag{2.21}$$

Un événement quelconque de cet univers peut s'écrire :

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{k} A_i \right) = \bigcup_{i=1}^{k} (B \cap A_i) = \bigcup_{i=1}^{k} (A_i \cap B) \tag{2.22}$$

Ainsi, la probabilité de réalisation de cet événement est :

$$p(B) = p\left[\bigcup_{i=1}^{k}(A_i \cap B)\right] = \sum_{i=1}^{k} p(A_i \cap B) = \sum_{i=1}^{k} p(A_i)\, p(B / A_i) \qquad (2.23)$$

C'est le **théorème des probabilités totales**. Considérons $p(A_i)$ la probabilité a priori. Dans ce cas, la probabilité conditionnelle de réalisation de l'événement B sachant que l'événement A_i est réalisé (probabilité a posteriori) vaut [13] :

$$p(A_i / B) = \frac{p(A_i \cap B)}{p(B)} = \frac{p(A_i)\, p(B / A_i)}{\sum_{i=1}^{k} p(A_i)\, p(B / A_i)} \qquad (2.24)$$

Telle est la formule de Bayes appelée communément théorème de Bayes. Ce théorème permet de calculer la probabilité a posteriori connaissant la probabilité a priori. Il trouve des applications pratiques notamment en faisant une inférence bayesienne, le calcul des probabilités a posteriori d'être malade sachant que le résultat du test est négatif en épidémiologie, … La probabilité a posteriori $p(A_i / B)$ se calcule à partir des probabilités a priori $p(A_i)$, de la vraisemblance $p(B / A_i)$ et de la distribution marginale $\sum_{i=1}^{k} p(A_i)\, p(B / A_i)$. C'est cette formule qui est à la base de la théorie bayésienne.

En épidémiologie, par exemple, il est courant d'utiliser une table de contingence qui croise le résultat du test de dépistage (positif ou négatif) et la maladie (malade ou pas malade) comme le montre le tableau 5 ci-après.

Tableau 5 : Table de contingence

Maladie	M+	M-	Total
T+	a	b	a+b
T-	c	d	c+d
Total	a+c	b+d	a+b+c+d

La valeur prédictive du résultat positif (probabilité a posteriori d'être malade si le test est positif) est :

$$VPP = p(M^+ / T^+) = \frac{p(M^+ \cap T^+)}{p(T^+)} = \frac{a}{a+b} \qquad (2.25)$$

Sachant que :

$$T^+ = (T^+ \cap M^+) \cup (T^+ \cap M^-) \Rightarrow p(T^+) = \left[p(T^+ \cap M^+) \cup (T^+ \cap M^-)\right]$$

il vient :

$$p\left(T^{+}\right)=p\left(T^{+}\cap M^{+}\right)+p\left(T^{+}\cap M^{-}\right)-p\overbrace{\left(\underbrace{T^{+}\cap\overbrace{M^{+}\cap M^{-}}^{o}}_{0}\right)}^{o}$$

$$=p\left(T^{+}/M^{+}\right).P\left(M^{+}\right)+p\left(T^{+}/M^{-}\right)p\left(M^{-}\right)$$

La VPP vaut donc [14] :

$$VPP=p\left(M^{+}/T^{+}\right)=\frac{p\left(T^{+}/M^{+}\right)p\left(M^{+}\right)}{p\left(T^{+}/M^{+}\right)p\left(M^{+}\right)+p\left(T^{+}/M^{-}\right)p\left(M^{-}\right)}$$

Par le même raisonnement, la valeur prédictive du résultat négatif (probabilité a posteriori d'être non malade si le test est négatif) est :

$$VPN=p\left(M^{-}/T^{-}\right)=\frac{p\left(M^{-}\cap T^{-}\right)}{p\left(T^{-}\right)}=\frac{p\left(T^{-}/M^{-}\right)p\left(M^{-}\right)}{p\left(T^{-}/M^{+}\right)p\left(M^{+}\right)+p\left(T^{-}/M^{-}\right)p\left(M^{-}\right)}=\frac{d}{c+d} \qquad (2.26)$$

2.7. Applications de la formule de Bayes

2.7.1. En médecine

En médecine, la formule de Bayes permet d'examiner les résultats d'un test diagnostic ou des examens des maladies. Considérons que la prévalence d'une maladie M dans une population donnée est de 25 %. Supposons, en outre, que la probabilité d'avoir un résultat négatif pour le test si l'individu n'est pas porteur de la maladie M soit de 0,9 et que la probabilité d'avoir un résultat positif si l'individu est porteur de la maladie M soit de 0,8. Dans ce cas, la probabilité pour qu'un individu pris au hasard dans cette population soit porteur de la maladie sachant que le test est positif vaut [10] :

Solution :

La probabilité conditionnelle permet, par exemple, d'évaluer les résultats d'un test de dépistage et des examens des maladies. Par exemple, le tableau suivant montre les résultats pour le diagnostic de la carie dentaire obtenus à partir d'une radiographie et un diagnostic basé sur l'examen dentaire clinique.

	M+	M-	Total
T+	400	10	410
T-	100	890	990
	500	900	1400

La question qui se pose est de déterminer la probabilité pour qu'un individu pris au hasard dans cette population soit malade sachant que le résultat du test est positif. Cela porte le nom de valeur prédictive positive ou valeur prédictive du résultat positif (VPP).

Le théorème de Bayes donne :

$$\begin{aligned} VPP = p\left(M^+ \mid T^+\right) &= \frac{p\left(M^+ \cap T^+\right)}{p\left(T^+\right)} \\ &= \frac{p\left(M^+\right) p\left(T^+ \mid M^+\right)}{p\left(M^+\right) p\left(T^+ \mid M^+\right) + p\left(M^-\right) p\left(T^+ \mid M^-\right)} \\ &= \frac{p\left(M^+\right) p\left(T^+ \mid M^+\right)}{p\left(M^+\right) p\left(T^+ \mid M^+\right) + p\left(M^-\right) p\left(T^+ \mid M^-\right)} \end{aligned} \quad (2.27)$$

$$P(T^+ \mid M^+) = \frac{P\left(M^+ \mid T^+\right) P\left(M^+\right)}{P\left(M^+ \mid T^+\right) P\left(M^+\right) + P\left(M^- \mid T^+\right) P\left(M^-\right)} = \frac{0{,}8 \times 0{,}25}{0{,}8 \times 0{,}25 + 0{,}2 \times 0{,}75} = 0{,}57$$

2.7.2. En sciences de l'éducation

Considérons le tableau 6 ci-après qui donne la répartition de 400 étudiants_d'une école selon le niveau d'études et l'option étudiée.

Tableau 6 : Répartition de 400 étudiants_d'une école selon le niveau d'études et l'option

Section	Mathématiques	Biologie-Chimie	Total
Classe			
1ère année	80	60	140
2ème année	75	25	100
3ème année	50	40	90
4ème année	45	25	70
Total	250	150	400

Un étudiant est choisi au hasard et il est inscrit en Mathématiques. La question qui se pose est de calculer la probabilité qu'il soit inscrit en 1ère année, en 2ème année, en 3ème année, en 4ème année ?

Solution :

Soient les événements :

N_1 : « Être inscrit en 1ère année »

N_2 : « Être inscrit en 2ème année »

N_3 : « Être inscrit en 3ème année »

N_4 : « Être inscrit en 4ème année »

M : « Être inscrit en Mathématiques »

B : « Être inscrit en Biologie-Chimie »

Les probabilités suivantes sont obtenues à partir du tableau des données :

$$P(N_1)=\frac{140}{400};P(N_2)=\frac{100}{400};P(N_3)=\frac{90}{400};P(N_4)=\frac{70}{400};P(M)=\frac{250}{400};P(B)=\frac{150}{400}$$

La probabilité d'être inscrit en 1ère année sachant que l'étudiant est en Mathématiques vaut :

$$P(N_1/M)=\frac{P(N_1\cap M)}{P(M)}$$

Or, si l'étudiant est inscrit en Mathématiques, cela veut dire qu'il est inscrit en Mathématiques et il est en 1ère année, ou bien il est inscrit en Mathématiques et il est en 2ème année, ou bien il est inscrit en Mathématiques et il est en 3ème année ou bien il est inscrit en Mathématiques et il est en 4ème année :

$$M=(M\cap N_1)\cup(M\cap N_2)\cup(M\cap N_3)\cup(M\cap N_4)$$

Donc :

$$\begin{aligned}p(M)&=p\left[(M\cap N_1)\cup(M\cap N_2)\cup(M\cap N_3)\cup(M\cap N_4)\right]\\&=p(M\cap N_1)+p(M\cap N_2)+p(M\cap N_3)+p(M\cap N_4)\\&=p(N_1)P(M/N_1)+p(N_2)P(M/N_2)+p(N_3)p(M/N_3)+p(N_4)p(M/N_4)\end{aligned}$$

Les probabilités qu'il soit inscrit en 1ère année sachant qu'il est en Mathématiques vaut :

$$\begin{aligned}p(N_1/M)&=\frac{p(N_1\cap M)}{p(M)}\\&=\frac{p(N_1)p(M/N_1)}{p(M)}\\&=\frac{p(N_1)p(M/N_1)}{p(N_1)p(M/N_1)+p(N_2)p(M/N_2)+p(N_3)P(M/N_3)+p(N_4)P(M/N_4)}\end{aligned}$$

Le remplacement donne :

$$\begin{aligned}p(N_1/M)&=\frac{\frac{140}{400}\times\frac{80}{140}}{\frac{140}{400}\times\frac{80}{140}+\frac{100}{400}\times\frac{75}{100}+\frac{90}{400}\times\frac{50}{90}+\frac{70}{400}\times\frac{45}{70}}\\&=\frac{\frac{80}{400}}{\frac{80}{400}+\frac{75}{400}+\frac{50}{400}+\frac{45}{400}}=\frac{80}{250}=0,32\end{aligned}$$

Les probabilités qu'il soit inscrit en 2ème année, en 3ème année, en 4ème année sont respectivement :

$$p(N_2/M) = \frac{p(N_2 \cap M)}{p(M)} = \frac{p(N_2)\,p(M/N_2)}{p(M)} = 0,30$$

$$p(N_3/M) = \frac{p(N_3 \cap M)}{p(M)} = \frac{p(N_3)\,p(M/N_3)}{p(M)} = 0,20$$

$$p(N_4/M) = \frac{p(N_4 \cap M)}{p(M)} = \frac{p(N_4)\,p(M/N_4)}{p(M)} = 0,18$$

2.8. Variables aléatoires

Une variable aléatoire X est une fonction qui, à un élément de l'univers (ensemble des éventualités), associe une valeur [15] :

$$\begin{aligned} X : \Omega &\to \mathbb{R} \\ \omega &\mapsto X(\omega) \end{aligned}$$

telle que $X^{-1}(\omega)$ est un événement. Les valeurs prises par une variable aléatoire sont appelées de réalisations de cette variable aléatoire et sont notées *x*. Ainsi, pour un événement donnée *A*, au lieu de raisonner en termes ensemblistes, le calcul des probabilités se fera en utilisant un raisonnement basé sur la notion de variable aléatoire [15] :

$$p(A) = p(X = x)$$

La variable *X* est appelée variable aléatoire réelle. Les valeurs prises par la variable aléatoire, appelées aussi réalisations de cette variable sont notées *x*.

2.8.1. Variable aléatoire continue

Si $X(\Omega)$ est non dénombrable, la variable X est dite continue. Elle est susceptible de prendre n'importe quelle valeur dans l'intervalle des réalisations. Cela veut dire que $X(\omega) \in \mathbb{R}$ ou que $X(\Omega) \subset \mathbb{R}$.

Une variable aléatoire continue est caractérisée par une fonction densité de probabilité $f(x)$ qui est telle que :

$$\int_{\mathbb{R}} f(x)\,dx = 1 \qquad (2.28)$$

Pour une variable continue, la probabilité qu'elle prenne une valeur donnée vaut :

$$p(X = x) \simeq 0 \qquad (2.29)$$

L'intégrale de la fonction de densité des probabilités donne une fonction de répartition [16] :

$$F(x) = p(X \leq x) = \int_{-\infty}^{x} f(t)\,dt \tag{2.30}$$

Cette fonction de répartition vérifie les propriétés :

$$\begin{aligned} &\lim_{x \to -\infty} F(x) = 0 \\ &\lim_{x \to +\infty} F(x) = 1 \end{aligned} \tag{2.31}$$

La fonction de répartition est :

- croissante ;
- constante sur l'intervalle $[x_i, x_{i+1}[$;
- nulle sur l'intervalle $]-\infty, x_1[$;
- égale à 1 sur l'intervalle $]x_n, +\infty[$.

L'espérance mathématique (ou moyenne espérée) d'une variable aléatoire continue se calcule comme suit :

$$E(X) = \int_{-\infty}^{+\infty} x\, f(x) \tag{2.32}$$

Par définition, la variance est

$$Var(X) = E\left\{\left[X - E(X)\right]^2\right\} \tag{2.33}$$

Le développement de cette fonction relation donne :

$$Var(X) = E(X^2) - E^2(X) \tag{2.34}$$

sachant que $E(X^2) = \int_{-\infty}^{+\infty} x^2\, f(x)$.

Exercice :

Soit X une variable aléatoire dont la fonction densité de probabilité est donnée par :

$$f(x) = \begin{cases} \dfrac{a}{x}\ si -e \leq x < -1 \\ x + 1 - a\ si -1 \leq x < 0 \\ 0\ sinon \end{cases}$$

Calculez la valeur de a pou que f soit une densité de probabilité.

Calculez sa fonction de répartition.

Solution :

La fonction f est une densité de probabilité si et seulement si :

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1 \Leftrightarrow \int_{-\infty}^{-e} 0dt + a\int_{-e}^{-1}\frac{dx}{x} + \int_{-e}^{-1}(x+1-a)\,dx + \int_{-e}^{-1} 0dx = 1$$

$$\Leftrightarrow 0 + a\left[ln|x|\right]_{-e}^{-1} + \left[\frac{x^2}{2} + x - ax\right]_{-1}^{0} + 0 = 1$$

$$\Leftrightarrow -a - \frac{1}{2} + 1 - a = 1$$

$$\Leftrightarrow -2a = \frac{1}{2}$$

$$\Leftrightarrow a = -\frac{1}{4}$$

La fonction densité s'écrit alors :

$$f(x) = \begin{cases} -\dfrac{1}{4x} \; si -e \le x < -1 \\ x + \dfrac{5}{4} \; si -1 \le x < 0 \\ 0 \; sinon \end{cases}$$

Si $x < -e$, la fonction de répartition de X est :

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0dt = 0$$

Si $-e \le x < -1$, la fonction de répartition de X est :

$$F(x) = 0 + \int_{-e}^{x} f(t)\,dt = -\frac{1}{4}\int_{-e}^{x}\frac{dt}{t} = -\frac{1}{4}\left[ln|t|\right]_{-e}^{x} = -\frac{1}{4}ln\frac{|x|}{e}$$

Si $-1 \le x < 0$, la fonction de répartition de X est :

$$F(x) = \frac{1}{4} + \int_{-1}^{x} f(t)\,dt = \frac{1}{4} + \int_{-1}^{x}\left(t + \frac{5}{4}\right)dt = \frac{1}{4} + \left[\frac{t^2}{2} + \frac{5}{4}t\right]_{-1}^{x} = \frac{1}{4} + \frac{x^2}{2} + \frac{5x}{4} - \frac{1}{2} + \frac{5}{4} = \frac{x^2}{2} + \frac{5x}{4} + 1$$

Si $x \ge 0$, la fonction de répartition de X est :

$$F(x) = 1 + \int_{0}^{x} f(t)\,dt = 1 + \int_{0}^{x} 0dt = 1$$

La fonction de répartition de *X* est donc :

$$F(x)=\begin{cases} 0 \; si \; x<-e \\ \dfrac{x^2}{2}+\dfrac{5x}{4}+1 \; si -e\le x<-1 \\ -\dfrac{1}{4}ln\dfrac{|x|}{e} \; si -1\le x<0 \\ 1 \; si \; x\ge 0 \end{cases}$$

2.8.2. Variable aléatoire discrète

Si, par contre, $X(\Omega)$ est dénombrable, la variable *X* est dite discrète ou discontinue. Cela veut dire que $X(\omega)\in \mathbb{N}$. La variable *X* ne prend que des valeurs isolées dans l'intervalle des réalisations. Pour une variable aléatoire discrète, sa loi de probabilité [15] est un tableau dont la première colonne contient les valeurs x_i prises par cette variable et la deuxième colonne les probabilités respectives $p(X=x_i)$ des événements élémentaires. Il faut noter que la somme de ces probabilités donne 1, dans la mesure où ce sont des événements élémentaires. Pour une variable aléatoire continue, sa fonction de répartition s'obtient en intégrant sur le domaine la fonction densité de cette variable. Par contre, pour une variable aléatoire discrète, la fonction de répartition s'obtient en faisant le cumul des probabilités à l'aide de la relation [17] :

$$F(x_i)=p(X\le x_i) \tag{2.35}$$

Les valeurs de cette fonction en chaque point se calculent comme suit :

$$\begin{aligned} &F(x_1)=p(X\le x_1)\\ &F(x_2)=p(X=x_1)+p(X=x_2)\\ &F(x_3)=p(X=x_1)+p(X=x_2)+p(X=x_3)\\ &\vdots\\ &F(x_n)=p(X=x_1)+p(X=x_2)+p(X=x_3)+...+p(X=x_n) \end{aligned}$$

Donner la densité de probabilité ou la fonction de répartition revient au même.

Sachant que :

$$F(x_j)=p(X=x_0)+p(X=x_1)+p(X=x_2)+...+p(X=x_{j-2})+p(X=x_{j-1})$$

il vient :

$$\begin{aligned} &F(x_j)=F(x_{j-1})+p(X=x_{j-1})\\ &p(X=x_{j-1})=F(x_j)-F(x_{j-1}) \end{aligned}$$

et finalement :

$$p\left(X = x_j\right) = F\left(x_{j+1}\right) - F\left(x_j\right)$$

Soit X la variable aléatoire représentant le nombre de points amené à chaque jeu d'un dé. L'ensemble des valeurs possibles est $X\left(\Omega\right) = \{1,2,3,4,5,6\}$. Chaque face du dé a la même chance d'apparaître, c'est l'équiprobabilité :

$$\begin{aligned} p\left(X = 1\right) &= \frac{1}{6} \\ p\left(X = 2\right) &= \frac{1}{6} \\ p\left(X = 3\right) &= \frac{1}{6} \\ p\left(X = 4\right) &= \frac{1}{6} \\ p\left(X = 5\right) &= \frac{1}{6} \\ p\left(X = 6\right) &= \frac{1}{6} \end{aligned}.$$

La fonction de répartition de X est :

$$F\left(x\right) = \begin{cases} 0 \text{ si } x < 1 \\ \frac{1}{6} \text{ si } 1 \le x < 2 \\ \frac{1}{3} \text{ si } 2 \le x < 3 \\ \frac{1}{2} \text{ si } 3 \le x < 4 \\ \frac{2}{3} \text{ si } 4 \le x < 5 \\ \frac{5}{6} \text{ si } 5 \le x < 6 \\ 1 \text{ si } x \ge 6 \end{cases}$$

Les probabilités se reconstituent en utilisant la relation :

$$p_i = p\left(x_i\right) = p\left(X = x_i\right) = F\left(x_i\right) - F\left(x_{i-1}\right)$$

L'espérance mathématique (ou moyenne espérée) d'une variable aléatoire discrète (ou discontinue) se calcule comme suit [17] :

$$E(X)=\sum_{k=1}^{n} p_k x_k \tag{2.36}$$

Par définition, la variance est :

$$Var(X)=\sum_{k=1}^{n} p_k \left[x_k - E(X)\right]^2 \tag{2.37}$$

Le développement de cette fonction relation donne :

$$Var(X)=\sum_{k=1}^{n} p_k x_k^2 - \left[\sum_{k=1}^{n} p_k x_k\right]^2 = E\left(X^2\right)-E^2(X) \tag{2.38}$$

Comme en statistique, si la variable X prend des valeurs $x_1, x_2, \ldots, x_n$, il est à noter que :

$$x_1 \leq E(X) \leq x_n \tag{2.39}$$

Les propriétés de l'espérance mathématique en tant qu'opérateur sont :

P₁ : L'espérance mathématique d'une constante est cette constante même.

$$E(a)=a, \forall a \in \mathbb{R} \tag{2.40}$$

P₂ : L'espérance mathématique d'une variable multipliée par une constante est égale à l'espérance mathématique de cette variable multipliée par cette constante [17].

$$E(aX)=aE(X), \forall a \in \mathbb{R} \tag{2.41}$$

P₃ : Lorsqu'une constance est ajoutée à une variable aléatoire, alors sont espérance mathématique est égale à l'espérance mathématique de cette variable augmentée de cette constante.

$$E(X+a)=E(X)+a, \forall a \in \mathbb{R} \tag{2.42}$$

P₄ : L'espérance est d'un produit d'une variable aléatoire multipliée par une constante, puis augmentée d'une autre constante est également au produit de l'espérance mathématique par la première constante augmentée de la deuxième constante.

$$E(aX+b)=aE(X)+b, \forall a,b \in \mathbb{R} \tag{2.43}$$

P₅ : Si X et Y sont deux variables aléatoires qui admettent une espérance, alors l'espérance d'une somme est la somme des espérances. L'espérance mathématique est donc un opérateur linéaire.

Cela montre que l'espérance mathématique d'une transformation linéaire est égale à la transformation linéaire de l'espérance mathématique :

$$E(X+Y)=E(X)+E(Y) \tag{2.44}$$

Si les variables aléatoires X et Y sont continues, alors :

$$\begin{aligned}
E(X+Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)\,f(x,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\,dx\,dy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty}x\left(\int_{-\infty}^{+\infty}f(x,y)\,dy\right)dx + \int_{-\infty}^{+\infty}y\left(\int_{-\infty}^{+\infty}f(x,y)\,dx\right)dy \\
&= \int_{-\infty}^{+\infty}x\left(\underbrace{\int_{-\infty}^{+\infty}f(x,y)\,dy}_{f(x)}\right)dx + \int_{-\infty}^{+\infty}y\left(\underbrace{\int_{-\infty}^{+\infty}f(x,y)\,dx}_{f(y)}\right)dy \\
&= \int_{-\infty}^{+\infty}xf(x)\,dx + \int_{-\infty}^{+\infty}yf(y)\,dy \\
&= E(X)+E(Y)
\end{aligned}$$

Comme conséquence, l'espérance d'une différence est la différence des espérances :

$$\begin{aligned}
E(X-Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-y)\,f(x,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\,dx\,dy - \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty}x\left(\int_{-\infty}^{+\infty}f(x,y)\,dy\right)dx - \int_{-\infty}^{+\infty}y\left(\int_{-\infty}^{+\infty}f(x,y)\,dx\right)dy \\
&= \int_{-\infty}^{+\infty}x\left(\underbrace{\int_{-\infty}^{+\infty}f(x,y)\,dy}_{f(x)}\right)dx - \int_{-\infty}^{+\infty}y\left(\underbrace{\int_{-\infty}^{+\infty}f(x,y)\,dx}_{f(y)}\right)dy \\
&= \int_{-\infty}^{+\infty}xf(x)\,dx - \int_{-\infty}^{+\infty}yf(y)\,dy \\
&= E(X)-E(Y)
\end{aligned}$$

De plus, si les variables aléatoires X et Y sont indépendantes, alors l'espérance d'un produit est égale au produit des espérances mathématiques [17] :

$$\begin{aligned}
E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(xy)\,f(x,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\,dx\,dy \times \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty} x\left(\int_{-\infty}^{+\infty} f(x,y)\,dy\right)dx \times \int_{-\infty}^{+\infty} y\left(\int_{-\infty}^{+\infty} f(x,y)\,dx\right)dy \\
&= \int_{-\infty}^{+\infty} x\left(\underbrace{\int_{-\infty}^{+\infty} f(x,y)\,dy}_{f(x)}\right)dx \times \int_{-\infty}^{+\infty} y\left(\underbrace{\int_{-\infty}^{+\infty} f(x,y)\,dx}_{f(y)}\right)dy \\
&= \int_{-\infty}^{+\infty} xf(x)\,dx \times \int_{-\infty}^{+\infty} yf(y)\,dy \\
&= E(X)\times E(Y)
\end{aligned}$$

P$_6$: L'espérance mathématique est un operateur linéaire :

$$E(\lambda X+\mu Y)=\lambda E(X)+\lambda E(Y);\forall\lambda,\mu\in\mathbb{R} \tag{2.45}$$

P$_7$: De plus, si $g(x)$ est une fonction continue quelconque alors :

$$E\left[g(X)\right]=\sum_{i=1}^{n} p_i g(X_i) \tag{2.46}$$

Soit X une variable aléatoire et Y son image à travers la fonction g continue, c'est-à-dire :

$$Y=g(X)$$

Dans ce cas, Y est aussi une variable aléatoire qui est telle que :

$$p(Y=y)=\sum_{x|g(x)=y} p(X=x)$$

Dans le cas discret, l'image de Y est :

$$E(Y)=E\left[g(X)\right]=\sum_{i=1}^{n} g(x_i)\,f(x_i)$$

Si, par contre, la variable aléatoire X est continue avec $f(x)$ comme fonction de probabilité, alors :

$$E\left[g(X)\right]=\int_{-\infty}^{+\infty} g(x)\,f(x)\,dx$$

Dans le cas bidimensionnel, c'est-à-dire lorsque la densité de probabilité jointe des variables aléatoires X et Y est $f(x, y)$, cela devient :

$$E\left[g(X,Y)\right]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x, y)\, f(x, y)\, dx\, dy$$

Les propriétés de la variance sont :

P₈ : La variance d'une constante est nulle[17].

$$Var(a)=0, \forall a \in \mathbb{R} \tag{2.47}$$

P₉ : La variance est toujours positive.

$$Var(X)\geq 0 \tag{2.48}$$

Preuve : La variance est la moyenne des carrés des écarts à la moyenne ou la somme des carrés des écarts à la moyenne divisée par le nombre d'observations (ou de réalisations en probabilité). De ce fait, elle est positive.

$$\begin{aligned}
&Var(X)=0\\
&\Leftrightarrow \sum_{i=1}^{n} p_i\left[x_i-E(X)\right]^2=0\\
&\Leftrightarrow p_i\left[x_i-E(X)\right]^2=0\\
&\Leftrightarrow p_i \text{ ou} \left[x_i-E(X)\right]^2=0\\
&\Leftrightarrow x_i-E(X)=0\\
&\Leftrightarrow E(X)=x_i
\end{aligned}$$

Dans ce cas, la variable aléatoire X est dite variable certaine ou variable de Dirac avec une probabilité égale à 1.

P₁₀ : La variance est invariante par translation ou par changement d'origine [17].

$$Var(X+a)=Var(X), \forall a \in \mathbb{R} \tag{2.49}$$

P₁₁ : La variance n'est pas invariante par changement d'échelle. Autrement dit, tout changement d'échelle modifie la variance :

$$Var(aX)=a^2 Var(X), \forall a \in \mathbb{R} \tag{2.50}$$

De fil en aiguille, la propriété 12 est une conséquence des deux propriétés précédentes.

P$_{12}$: La variance n'est pas invariante par changement d'échelle.

$$Var\left(aX+b\right)=a^2Var\left(X\right),\forall a,b\in \mathbb{R} \tag{2.51}$$

P$_{13}$: La variance d'une somme n'est pas égale à la somme des variances des deux variables respectives comme pour l'espérance mathématique. Cela est vrai uniquement lorsque les variables aléatoires sont indépendantes, ce qui veut dire que la covariance entre est deux variables est nulle :

$$Var\left(X\pm Y\right)=Var\left(X\right)+Var\left(X\right)\pm 2Cov\left(X,Y\right) \tag{2.52}$$

Exemple 1 :

Considérons la variable aléatoire *X* qui représente le nombre de succès lorsqu'une pièce de monnaie est lancée, sachant que le succès consiste en l'obtention de la face PILE et l'échec en l'obtention de la face FACE. La variable *X* prend deux valeurs possibles 0 et 1 avec les probabilités respectives égales à 0,5. C'est l'équiprobabilité des faces.

Ainsi, les moments (espérance mathématique et variance s'obtiennent comme suit :

$$E\left(X\right)=\sum_{k=0}^{1}p_kx_k=\frac{1}{2}\times 0+\frac{1}{2}\times 1=\frac{1}{2}$$

$$\Rightarrow E\left(X^2\right)=\sum_{k=0}^{1}p_kx_k^2=\frac{1}{2}\times 0^2+\frac{1}{2}\times 1^2=\frac{1}{2}$$

$$Var\left(X\right)=E\left(X^2\right)-E^2\left(X\right)=\frac{1}{2}-\left(\frac{1}{2}\right)^2=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$$

Exemple 2 :

Soient deux variables aléatoires *X* et *Y* dont les réalisations et les probabilités correspondantes sont consignées dans le tableau ci-après. La question qui se pose est de savoir la variable aléatoire qui a une grande dispersion.

	Variable *X*			Variable *Y*		
Réalisations	2	4	6	-4	3	33
Probabilités	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

Les moments (espérance mathématique et variance) de *X* sont :

$$E\left(X\right)=\sum_{k=1}^{3}p_kx_k=2\times\frac{1}{4}+4\times\frac{1}{4}+6\times\frac{1}{2}=\frac{2+4+12}{4}=\frac{18}{4}=4,5$$

$$Var\left(X\right)=\sum_{k=0}^{1}p_k\left[x_k-E\left(X\right)\right]^2=\frac{1}{4}\times\left(2-4,5\right)^2+\frac{1}{4}\times\left(4-4,5\right)^2+\frac{1}{2}\times\left(6-4,5\right)^2=2,75$$

Les moments (espérance mathématique et variance) de *Y* sont :

$$E(Y)=\sum_{k=1}^{3} p_k y_k = -4\times\frac{1}{2}+3\times\frac{1}{3}+33\times\frac{1}{6}=-2+1+\frac{11}{2}=\frac{11-2}{2}=4,5$$

$$Var(Y)=\sum_{k=0}^{1} p_k\left[y_k-E(Y)\right]^2=\frac{1}{2}\times(-4-4,5)^2+\frac{1}{3}\times(3-4,5)^2+\frac{1}{6}\times(33-4,5)^2=172,25$$

C'est donc la variable *Y* qui a une grande dispersion car elle est de plus grande variance : $Var(Y)>Var(X)$.

Pour illustration dans la vie courante, considérons la variable aléatoire *X* qui caractérise le nombre de garçons dans une famille de 4 enfants est une variable aléatoire discrète. Les valeurs qu'elle est susceptible de prendre sont des valeurs entières 0, 1, 2, 3 et 4.

Ainsi, les probabilités des événements élémentaires sont :

$$p(X=0)=C_4^0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4=P(FFFF)=0,0625\simeq 0,06$$

$$p(X=1)=C_4^1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3=4P(GFFF)=0,25$$

$$p(X=2)=C_4^2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2=6P(GGFF)=0,375\simeq 0,38$$

$$p(X=3)=C_4^3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1=4P(GGGF)=0,25$$

$$p(X=4)=C_4^4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0=P(GGGG)=0,0625\simeq 0,06$$

La loi de probabilité de cette variable est représentée dans le tableau 7 ci-après (les deux premières colonnes).

Tableau 7 : Loi de probabilité

x_i	p_i	p_i cumulée	$p_i\ x_i$
0	0,0625	0,0625	0,00
1	0,25	0,3125	0,25
2	0,375	0,6900	0,75
3	0,25	0,9375	0,75
4	0,0625	1,00	0,25
Σ	1,00	/////////////////	2,00

Comme la variable est discrète, cette loi peut se présenter sous forme de diagramme en bâtons (Figure 4).

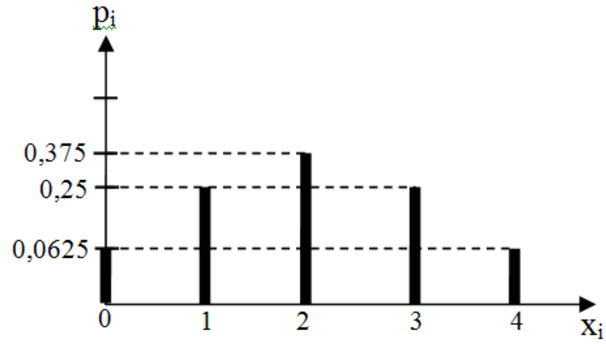


Figure 4 : Diagramme en bâtons
Source : Pr BARANKANIRA Emmanuel

La fonction de répartition de X se calcule comme suit :

$$F(x) = \begin{cases} 0 \text{ si } x \leq 0 \\ 0{,}0625 \text{ si } 0 < x \leq 1 \\ 0{,}3125 \text{ si } 1 < x \leq 2 \\ 0{,}690 \text{ si } 2 < x \leq 3 \\ 0{,}9375 \text{ si } 3 < x \leq 4 \\ 1 \text{ si } x > 4 \end{cases}$$

Cette fonction de répartition est représentée sous forme d'un diagramme en escalier de la figure 5.

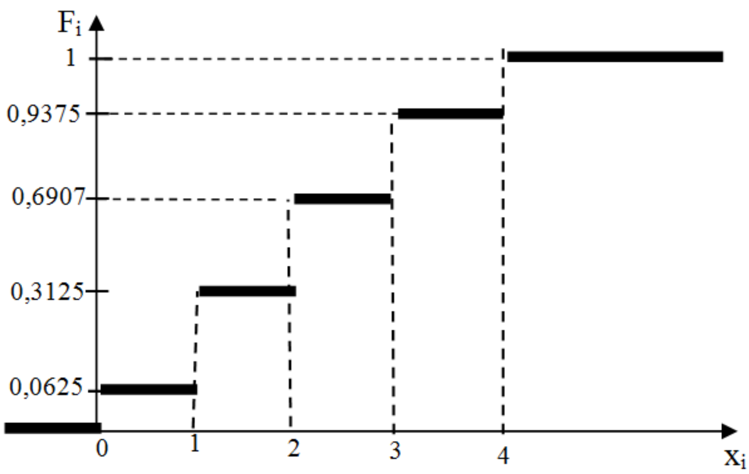


Figure 5 : Diagramme en escalier
Source : Pr BARANKANIRA Emmanuel

Après calculs :

$$E(X) = \sum_{k=1}^{n} p_k x_k = 2$$

Cela montre que, dans une famille de 4 enfants, le nombre espéré de garçons vaut 2.

La variance vaut :

$$\begin{aligned}Var(X) &= 0{,}0625\times(0-2)^2+0{,}25\times(1-2)^2+0{,}375\times(2-2)^2+0{,}25\times(3-1)^2+0{,}0625\times(4-2)^2\\&=4(0{,}0625)+0{,}25+0{,}25+4(0{,}0625)\\&=0{,}50+0{,}50\\&=1\end{aligned}$$

Il est à remarquer que les opérations de centrage et de réduction donnent :

$$E\left[X-E(X)\right]=E(X)-E\left[E(X)\right]=E(X)-E(X)=0$$

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

$$E\left[\frac{X-E(X)}{\sigma}\right]=\frac{1}{\sigma}E\left[X-E(X)\right]=\frac{1}{\sigma}\left[E(X)-E\left(E(X)\right)\right]=\frac{1}{\sigma}\left[E(X)-E(X)\right]=0$$

$$Var\left[\frac{X-E(X)}{\sigma}\right]=\frac{1}{\sigma^2}Var\left[X-E(X)\right]=\frac{1}{\sigma^2}Var(X)=\frac{\sigma^2}{\sigma^2}=1$$

Considérons, à présent, une variable aléatoire continue *X* définies par la fonction de densité de probabilité suivante :

$$f(x)=\begin{cases}1 \text{ si } 0\le x\le 1\\ 0 \text{ sinon}\end{cases}$$

Les moments de *X* sont :

$$E(X)=\int_0^1 xf(x)dx=\int_0^1 xdx=\left[\frac{x^2}{2}\right]_0^1=\frac{1}{2}$$

$$\Rightarrow E(X^2)=\int_0^1 x^2f(x)dx=\int_0^1 x^2dx=\left[\frac{x^3}{3}\right]_0^1=\frac{1}{3}$$

$$Var(X)=E(X^2)-E^2(X)=\frac{1}{3}-\frac{1}{4}=\frac{1}{12}$$

En outre, considérons l’exemple plus haut relatif à la variable continue *X* par morceaux et dont la densité est reprise ci-après [17] :

$$f(x)=\begin{cases}\frac{a}{x} \text{ si } -e\le x<-1\\ x+1-a \text{ si } -1\le x<0\\ 0 \text{ sinon}\end{cases}$$

Quelle est l’espérance mathématique de *X* ?

Solution :

L'espérance mathématique de X vaut :

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} xf(x)\,dx \\
&= \int_{-\infty}^{-e} 0dx - \frac{1}{4}\int_{-e}^{-1} dx + \int_{-1}^{0}\left(x^2 + \frac{5}{4}x\right)dx + \int_{0}^{+\infty} 0dx \\
&= 0 - \frac{1}{4}[x]_{-e}^{-1} + \left[\frac{x^3}{3} + \frac{5}{8}x^2\right]_{-1}^{0} + 0 \\
&= -\frac{1}{4}(-1+e) - \left(-\frac{1}{3} + \frac{5}{8}\right) \\
&= \frac{1}{4} - \frac{e}{4} + \frac{1}{3} - \frac{5}{8} \\
&= \frac{6-6e+8-15}{24} \\
&= \frac{-6e-1}{24} \\
&= \frac{-1+6e}{24}
\end{aligned}
$$

Le moment non centré d'ordre $r \in \mathbb{N}_0$ lorsqu'il existe s'écrit :

$$m_r(X) = E\left(X^r\right) = \int_{-\infty}^{+\infty} x^r f(x)\,dx$$

Le moment centré d'ordre $r \in \mathbb{N}_0$ lorsqu'il existe s'écrit :

$$\mu_r(X) = E\left\{\left[x - E(X)\right]^r\right\} = \int_{-\infty}^{+\infty}\left[x - E(X)\right]^r f(x)\,dx$$

Exercice :

Une poche contient 5 billets de 20F, 3 billets de 50F et 1 billet de 200F. Ces billets sont indiscernables au toucher. Trois billets sont tirés aléatoirement. Soit X la variable aléatoire représentent le montant de ces 3 billets.

Déterminez la loi de probabilité de X et calculez l'espérance mathématique, la variance et l'écart-type de X.

Solution :

Les billets sont :

5 billets de 20F ; 3 billets de 50F ;

1 billet de 200F

Montant (X)	20F	50F	200F
60F	3	0	0
90F	2	1	0
120F	1	2	0
150F	0	3	0
240F	2	0	1
270F	1	1	1
300F	0	2	1

Soit X la variable aléatoire représentant le montant. Les valeurs prises par cette variable sont : 60, 90, 120, 150, 240, 270 et 300. La composition en termes de billets est reprise dans le tableau ci-après. Les valeurs dans le tableau représentent le nombre de billets de chaque type.

La probabilité que le montant soit de 60F est :

$$p(X=60)=\frac{C_5^3}{C_9^3}=\frac{\frac{5!}{3!2!}}{\frac{9!}{3!6!}}=\frac{\frac{5(4)}{2}}{\frac{9(8)(7)}{6}}=\frac{10}{84}\simeq 0,12$$

La probabilité que le montant soit de 90F est :

$$p(X=90)=\frac{C_5^2C_3^1}{C_9^3}=\frac{\frac{5!}{2!3!}\times 3}{\frac{9!}{3!6!}}=\frac{\frac{5(4)(3)}{2}}{\frac{9(8)(7)}{6}}=\frac{30}{84}\simeq 0,36$$

La probabilité que le montant soit de 120F est :

$$p(X=120)=\frac{C_3^2C_5^1}{C_9^3}=\frac{\frac{3!}{2!1!}\times 5}{\frac{9!}{3!6!}}=\frac{\frac{6(5)}{2}}{\frac{9(8)(7)}{6}}=\frac{15}{84}\simeq 0,18$$

La probabilité que le montant soit de 150F est :

$$p(X=150)=\frac{C_3^3}{C_9^3}=\frac{1}{\frac{9!}{3!6!}}=\frac{1}{\frac{9(8)(7)}{6}}=\frac{1}{84}\simeq 0,01$$

La probabilité que le montant soit de 240F est :

$$p(X=150)=\frac{C_5^2C_1^1}{C_9^3}=\frac{\frac{5!}{2!3!}}{\frac{9!}{3!6!}}=\frac{\frac{5(4)}{2}}{\frac{9(8)(7)}{6}}=\frac{10}{84}\simeq 0,12$$

La probabilité que le montant soit de 270F est :

$$p(X=270)=\frac{C_5^1C_3^1C_1^1}{C_9^3}=\frac{5(3)}{\frac{9!}{3!6!}}=\frac{15}{\frac{9(8)(7)}{6}}=\frac{15}{84}\simeq 0,18$$

La probabilité que le montant soit de 300F est :

$$p(X=300)=\frac{C_1^1C_3^2}{C_9^3}=\frac{\frac{3!}{2!1!}}{\frac{9!}{3!6!}}=\frac{3}{\frac{9(8)(7)}{6}}=\frac{3}{84}\simeq 0,03$$

La loi de probabilité est :

x_i	60	90	120	150	240	270	300	Σ
p_i	0,12	0,36	0,18	0,01	0,12	0,18	0,03	1
p_ix_i	7,2	32,4	21,6	1,5	28,8	48,6	9	149,1
$p_ix_i^2$	432	2916	2592	225	6912	13122	2700	28899

L'espérance mathématique de *X* vaut :

$$E(X)=\sum_{i=1}^{n}p_ix_i=149,1\simeq 150$$

La variance *X* vaut:

$$Var(X)=E(X^2)-E^2(X)=\sum_{i=1}^{n}p_ix_i^2-(150)^2\simeq 29250-22500\simeq 6750$$

Son écart-type vaut :

$$\sigma=\sqrt{Var(X)}=82,16$$

2.8.3. Variables aléatoires indépendantes

Soient Ω un ensemble fini non vide muni d'une probabilité p, *X* et *Y* deux variables aléatoires définies sur Ω, $x_1,x_2,...,x_n$ les éléments de $X(\Omega)$ et $y_1,y_2,...,y_q$ les éléments de $Y(\Omega)$. La fonction numérique *p* définie sur $X(\Omega)\times Y(\Omega)$ par la relation $p_{ij}=p(x_i,y_j)=p(X^{-1}(x_i),Y^{-1}(y_j))=p(X=x_i,Y=y_j)$ porte le nom de loi jointe (ou conjointe) des variables aléatoires *X* et *Y* ou encore de loi de la variable aléatoire à deux dimensions *X* et *Y*.

Cette fonction vérifie les propriétés suivantes :

$$p(x_i,y_j)\geq 0$$

$$\sum_{i=1}^{n}\sum_{j=1}^{q}p(x_i,y_j)=1,\forall(x_i,y_j)\in X(\Omega)\times Y(\Omega)$$

La loi jointe p des variables aléatoires *X* et *Y* définies sur $X(\Omega)\times Y(\Omega)$ définit donc une probabilité. À partir de la loi conjointe, les lois marginales de la loi *p* sont définies.

Ainsi, la probabilité marginale des lignes s'écrit :

$$p_{i\bullet} = p(X = x_i) = \sum_{j=1}^{q} p(x_i, y_j)$$

$$p_{\bullet j} = p(Y = y_j) = \sum_{i=1}^{n} p(x_i, y_j)$$

Le tableau 8 montre les probabilités jointes et marginales variables *X* et *Y*.

Tableau 8 : Probabilités jointes et marginales variables *X* et *Y*

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_q	Loi marginale de *Y*
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	p_{1q}	$p_{1\bullet}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	p_{2q}	$p_{2\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	p_{iq}	$p_{i\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nj}	$\cdots$	p_{nq}	$p_{n\bullet}$
Loi marginale de *X*	$p_{\bullet 1}$	$p_{\bullet 2}$	$\cdots$	$p_{\bullet j}$	$\cdots$	$p_{\bullet q}$	$p_{\bullet\bullet}$

La loi marginale de *X* et de *Y* est représentée par le tableau 9 ci-après.

Tableau 9 : Loi marginale de *X* et de *Y*

Loi marginale de *X*		Loi marginale de *Y*	
X	$p_{i\bullet}$	**Y**	$p_{\bullet j}$
x_1	$p_{1\bullet}$	y_1	$p_{\bullet 1}$
x_2	$p_{2\bullet}$	y_2	$p_{\bullet 2}$
$\vdots$	$\cdots$	$\vdots$	$\cdots$
x_i	$p_{i\bullet}$	y_j	$p_{\bullet j}$
$\vdots$	$\cdots$	$\vdots$	$\cdots$
x_n	$p_{n\bullet}$	y_q	$p_{\bullet q}$
	$p_{\bullet\bullet} = 1$		$p_{\bullet\bullet} = 1$

En supposant que la probabilité que $X = x_i$ soit non nulle $\forall i \in [1,n]$, alors la probabilité conditionnelle de l'événement que $Y = y_j$ sachant $X = x_i$ est égale à :

$$p_{X=x_i}\left(Y = y_j\right) = \frac{p_{ij}}{p_{i\bullet}} \tag{2.53}$$

De même, si la probabilité que l'événement $Y = y_j$ est non nulle $\forall j \in [1,q]$, alors la probabilité conditionnelle de l'événement que $X = x_i$ sachant $Y = y_j$ est égale à :

$$p_{Y=y_j}\left(X = x_i\right) = \frac{p_{ij}}{p_{\bullet j}} \tag{2.54}$$

Deux variables aléatoires *X* et *Y* sont alors indépendantes si :

$$p_{j/i} = p_{X=x_i}\left(Y = y_j\right) = p\left(Y = y_j\right) = p_{\bullet j}$$

$$p_{i/j} = p_{Y=y_j}\left(X = x_i\right) = p\left(X = x_i\right) = p_{i\bullet} \tag{2.55}$$

La donnée de la loi conjointe de deux variables aléatoires *X* et *Y* détermine les lois marginales de *X* et *Y*. Réciproquement, la donnée de deux lois marginales de *X* et *Y* détermine la loi conjointe de deux variables aléatoires *X* et *Y* à condition que ces deux variables soient indépendantes.

Exemple :

Deux dés dont les faces sont équiprobables sont jetés. Les variables aléatoires *X* et *Y* désignent le nombre de points amenés par le premier dé et le plus grand nombre de points amenés par les deux dés respectivement.

Déterminer la loi conjointe des deux variables aléatoires *X* et *Y*. Monter que les deux variables aléatoires *X* et *Y* sont indépendantes. Déterminer deux probabilités conditionnelles.

Solution :

Jeter un dé deux fois revient à jeter deux dés en même temps. Les couples obtenus sont consignés dans le tableau 10.

Tableau 10 : Couples de points obtenus

2ème dé 1er dé	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2		(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3			(3,3)	(3,4)	(3,5)	(3,6)
4				(4,4)	(4,5)	(4,6)
5					(5,5)	(5,6)
6						(6,6)

Loi conjointe

2ème dé / 1er dé	1	2	3	4	5	6	Total
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
2		$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
3			$\frac{3}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
4				$\frac{4}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
5					$\frac{5}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
6						$\frac{6}{36}$	$\frac{1}{36}$
Total	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	$\frac{1}{6}$

Deux variables aléatoires *X* et *Y* sont indépendantes ssi $p_{ij} \neq p_{i\bullet} p_{\bullet j}, \forall i, j$

$$\left.\begin{array}{l} p_{22} = \dfrac{2}{36} \\ p_{2\bullet} p_{\bullet 2} = \dfrac{1}{6}\dfrac{3}{36} = \dfrac{1}{72} \end{array}\right\} \Rightarrow p_{22} \neq p_{2\bullet} p_{\bullet 2}$$

Les deux variables aléatoires *X* et *Y* ne sont pas indépendantes.

a) Deux probabilités conditionnelles sont :

$$p_{Y=2}\left(X = 2\right) = \frac{p_{22}}{p_{\bullet 2}} = \frac{\frac{2}{36}}{\frac{3}{36}} = \frac{2}{3}$$

$$p_{X=5}\left(Y = 1\right) = \frac{p_{52}}{p_{5\bullet}} = \frac{0}{\frac{1}{6}} = 0$$

2.9. Exercices d'entraînement

Exercice 1 :

Cent (100) personnes ont été interrogées en leur demandant pour lequel des partis politiques A, B et C elles vont voter. Les résultats sont regroupés dans le tableau ci-après.

Parti	A	B	C	Total
Homme	13	21	19	53
Femme	20	8	19	47
Total	33	29	38	100

Une personne au hasard est choisie dans ce groupe. Quelle est la probabilité qu'elle vote.

a) pour A ?
b) pour A sachant que c'est une femme ?
c) pour B ou C sachant que c'est une femme ?
d) pour B ou C sachant que c'est un homme ?
e) que ce soit une femme si la personne a voté pour le parti C ?
f) que ce soit un homme si la personne a voté pour C ?
g) pour A ou B ou C ?

Exercice 2 :

Trois urnes U_1, U_2 et U_3 dont les compositions sont indiquées dans le tableau 7 ci-après sont tirées au sort. Sachant qu'une boucle rouge a été obtenue, quelle est la probabilité qu'elle provienne dans l'urne U_2 ?

Couleur	Rouge	Blue	Verte	Total
U_1	3	4	1	8
U_2	1	2	3	6
U_3	4	3	2	9
Total	8	9	6	23

Exercice 3 :

Considérons 3 urnes identiques. La $1^{ère}$ contient 2 boucles blanches et une noire. La $2^{ème}$ contient 3 boucles blanches et une noire. La troisième contient 2 boules blanches et 2 noirs. Une urne est choisie au hasard et une boucle en est tirée. Trouver la probabilité de tirer une boucle blanche.

Chapitre 3. Lois usuelles discrètes

3.1. Loi de Dirac

Soit $a \in \mathbb{R}$, un point fixe. La loi Dirac, notée δ_a, est la loi de la variable aléatoire qui est constante et qui prend la valeur *a* quel que soit le résultat de l'épreuve [17] :

$$X(\omega) = a, \forall \omega \subset \Omega \Rightarrow p(X = a) = 1 \tag{3.1}$$

De ce fait, la loi de probabilité de X est donnée dans le tableau ci-après :

x_i	p_i	$p_i x_i$
a	1	a

La fonction de répartition de X est :

$$F(x) = \begin{cases} 0 \text{ si } x < a \\ 1 \text{ si } x \geq 0 \end{cases} \tag{3.2}$$

Les moments d'une variable de Dirac, appelée aussi variable certaine, sont :

$$E(X) = \sum_{k=1}^{n} p_k x_k = \sum_{i=1}^{n} a p_k = a \sum_{k=1}^{n} p_k = a \tag{3.3}$$

$$Var(X) = \sum_{k=1}^{n} p_k \left[x_k - E(X) \right]^2 = \sum_{k=1}^{n} p_k (a - a) = 0 \tag{3.4}$$

L'écart-type de *X* vaut alors :

$$\sigma_X = 0 \tag{3.5}$$

3.2. Loi uniforme discrète

Soit *n* un nombre naturel non nul. Une variable aléatoire *X* suit une loi uniforme lorsque toutes les modalités de cette variable ont une même probabilité égale à 1 divisé par le nombre *n* de réalisations de cette variable. Par exemple, lorsqu'un dé non pipé est jeté, l'ensemble des valeurs prises par la variable aléatoire qui désigne le nombre de points amené le dé est $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$. Chaque face a la même chance d'apparaître [17] :

$$p(X = k) = \frac{1}{n} \tag{3.6}$$

Autrement dit, l'application qui, à tout $k \in [1, n]$ associe la probabilité $p = \frac{1}{n}$ est la loi d'une variable aléatoire qui suit une loi uniforme discrète de paramètre *n*. Cette variable est appelée variable uniforme discrète. Cette variable se note : $X \sim U(n)$.

L'espérance mathématique et la variance d'une variable aléatoire qui suit une loi uniforme discrète de paramètre *n* sont (**à démontrer**) [17] :

$$E(X) = \frac{n+1}{2}$$
$$Var(X) = \frac{n^2-1}{12} \tag{3.7}$$

L'écart-type de *X* vaut alors :

$$\sigma_X = \frac{\sqrt{n^2-1}}{2\sqrt{3}} \tag{3.8}$$

Exemple :

Un dé non pipé est lancé. L'ensemble des valeurs de la variable aléatoire qui représente le nombre de points amenés par ce dé est $X(\Omega) = \{1,2,3,4,5,6\}$. Comme chaque face a la même chance d'apparaître, alors $p(X = k) = \frac{1}{6}$ $avec\ k = \overline{1,6}$.

3.3. Loi de Bernoulli

De façon générale, une variable aléatoire [17] :

$$X(\omega) = \begin{cases} 1\ si\ \omega \in A \\ 0\ si\ \omega \notin A \end{cases} \quad \text{avec } A \cup \overline{A} = \Omega \tag{3.9}$$

est une variable de Bernoulli. Dans ce cas, l'ensemble des valeurs possibles que cette variable aléatoire peut prendre est $X(\Omega) = \{0,1\}$. La variable aléatoire X suit une loi de Bernoulli de paramètre $p = p(A)$ et se note $X \to Bin(1, p)$ ou $X \sim Bin(1, p)$ avec p la probabilité de succès et $1 - p = q$ la probabilité d'échec. *Bin* réfère à la loi binomiale (voir plus loin) en tant que généralisation de la loi de Bernoulli et cette loi peut se noter simplement $X \sim Ber(p)$.

Considérons un jet d'une pièce d'une monnaie une seule fois. L'univers des possibles est $\Omega = \{P, F\}$. Le succès consiste en l'obtention de la face PILE et l'échec en l'obtention de la face FACE. Soit *X* désigne la variable aléatoire qui désigne le nombre de succès.

Cette variable est binaire et, de ce fait, elle prend les valeurs :

$$X = \begin{cases} 1\ si\ PILE \\ 0\ si\ FACE \end{cases}$$

soit

$$X = \begin{cases} 0 \text{ si échec} \\ 1 \text{ si succès} \end{cases}$$

Cette variable est appelée variable binaire ou variable de Bernoulli. Il s'agit d'un cas particulier de variable dichotomique, c'est-à-dire une variable qui prend deux modalités (malade ou sain, succès ou échec, présence ou absence, guérison ou rechute, décès ou survie, homme ou femme, etc). Posons p la probabilité de succès (probabilité d'obtenir la face PILE) et $q=1-p$ la probabilité d'échec (probabilité d'obtenir la face FACE).

Dans ce cas, au lieu de raisonner en termes d'événements, il est possible de raisonner en termes de variables aléatoires pour calculer les probabilités :

$$\begin{aligned} &p(\text{succès}) = p(X=1) = p \\ &p(\text{échec}) = p(X=0) = 1-p = q \end{aligned} \tag{3.10}$$

La variable aléatoire X suit une loi de Bernoulli de paramètre p, ce qui se note [17] :

$$X \sim Ber(p) \tag{3.11}$$

La probabilité qu'une variable aléatoire X qui suit une loi de Bernoulli de paramètre p prenne une valeur donnée k vaut :

$$p(X=k) = p^k(1-p)^{1-k} \quad \text{avec } k \in \{0,1\} \tag{3.12}$$

La loi de probabilité est :

x_i	p_i	F_i	$p_i x_i$
0	$1-p$	$1-p$	0
1	p	1	p
	1		P

L'espérance mathématique d'une variable aléatoire X qui suit une loi de Bernoulli de paramètre p sont respectivement (**à démontrer**) [17] :

$$\begin{aligned} &E(X) = p \\ &Var(X) = p(1-p) \end{aligned} \tag{3.13}$$

L'écart-type de la loi de Bernoulli est donc :

$$\sigma_X = \sqrt{p(1-p)} \tag{3.14}$$

Les coefficients de symétrie et d'aplatissement de la loi de Bernoulli sont respectivement :

$$\begin{aligned} \alpha_3 &= \frac{q-p}{\sqrt{pq}} = \frac{1-2p}{\sqrt{p(1-p)}} \\ \alpha_4 &= 3+\frac{1-6pq}{\sqrt{pq}} = 3+\frac{1-6p(1-p)}{\sqrt{p(1-p)}} \end{aligned} \tag{3.15}$$

Exemple :

Dans une population donnée, la prévalence d'une maladie M est de 10 %. Un individu est tiré aléatoirement dans cette population. Quelle est la probabilité qu'est soit porteuse de la maladie ?

Solution :

Considérons les événements :

Succès : « La personne est porteuse de la maladie M » => p(succès)=p=0,10

Echec : « La personne est saine » => p(échec)=q=1-p=0,90

C'est la loi de Bernoulli. Notons X la variable aléatoire qui représente le nombre de personnes malades. Si une seule personne est tirée, alors le nombre de personnes porteuses de la maladie M est soit égal à 0, soit égal à 1. La variable aléatoire X suit donc une loi de Bernoulli de paramètre p=0,10 : $X \sim Bern(0,10)$

La probabilité que la personne soit porteuse de la maladie vaut :

$$p(X=1) = p^1(1-p)^{1-1} = p = 0,10$$

3.4. Loi binomiale

La binomiale est une généralisation de la loi de Bernoulli. En effet, considérons n épreuves successives indépendantes au cours desquelles un événement A est réalisé ou pas pour chaque épreuve. Dans ce cas, il s'agit d'une suite de la forme $\underbrace{A\bar{A}AA\bar{A}A\bar{A}...AA\bar{A}A}_{n\ fois}$. À un événement élémentaire ω est associé un nombre $X(\omega)$ de réalisation de A. Autrement dit, si l'événement A se réalise, alors il y a succès et s'il ne se réalise pas, il y a échec. C'est la loi de Bernoulli. Comme l'expérience est répétée n fois, alors il s'agit d'une loi binomiale. La variable aléatoire X qui représente le nombre de succès est une variable qui suit une loi binomiale de paramètres n (le nombre de fois que l'expérience aléatoire de Bernoulli est répétée ou la taille de l'échantillon) et p (la probabilité de succès). Il est à noter que $p = p(A)$ et $X(\Omega) = \{0,1,2,...,n\}$. Cette variable est appelée variable binomiale et représente le nombre de succès réalisés au cours de cette suite donnée d'épreuves.

Elle se note [17] :

$$X \sim Bin(n,p) \tag{3.16}$$

La probabilité qu'une variable aléatoire X qui suit une loi binomiale de paramètres n et p prenne une valeur donnée k s'écrit [11] :

$$p(X=k)=C_n^k p^k (1-p)^{n-k} \ avec \ k \in \{0,1,\cdots,n\} \tag{3.17}$$

La figure 6 montre la distribution des probabilités calculées à l'aide de la loi binomiale de paramètres n=20 et p=0,50.

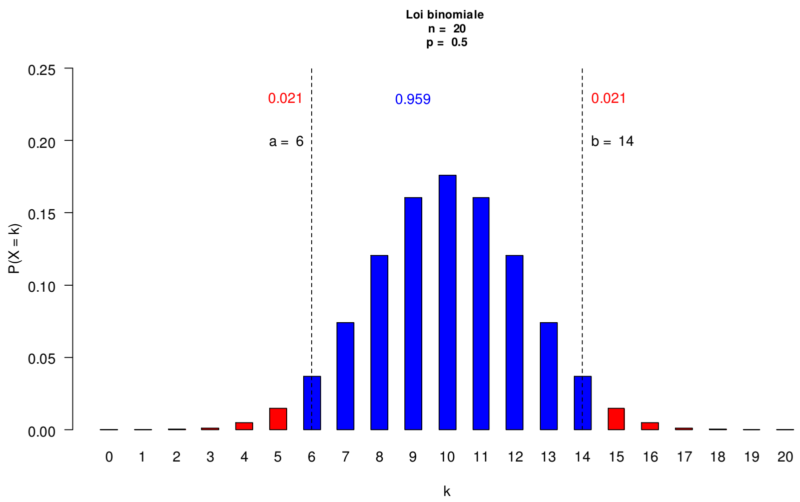


Figure 6 : Loi binomiale
Source : Pr BARANKANIRA Emmanuel

Les figures 7 et 8 montrent la fonction de répartition d'une variable binomiale de paramètres n=20 et p=0,50.

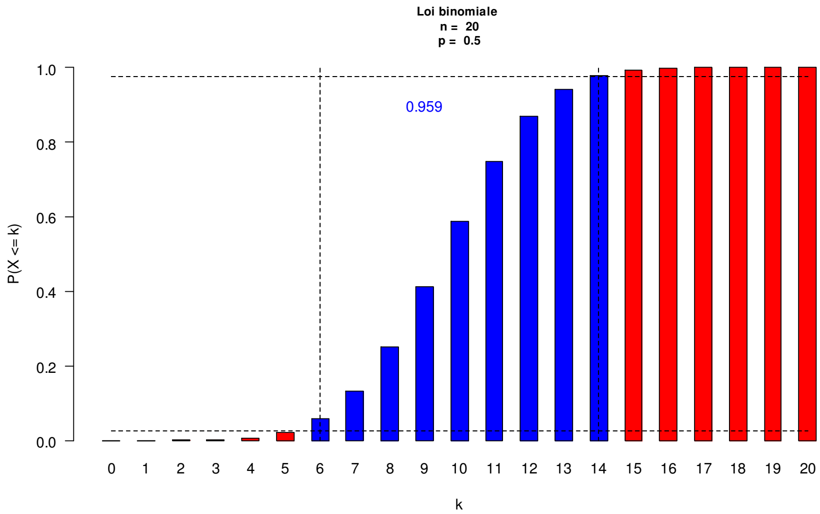


Figure 7 : Fonction de répartition d'une variable binomiale
Source : Pr BARANKANIRA Emmanuel

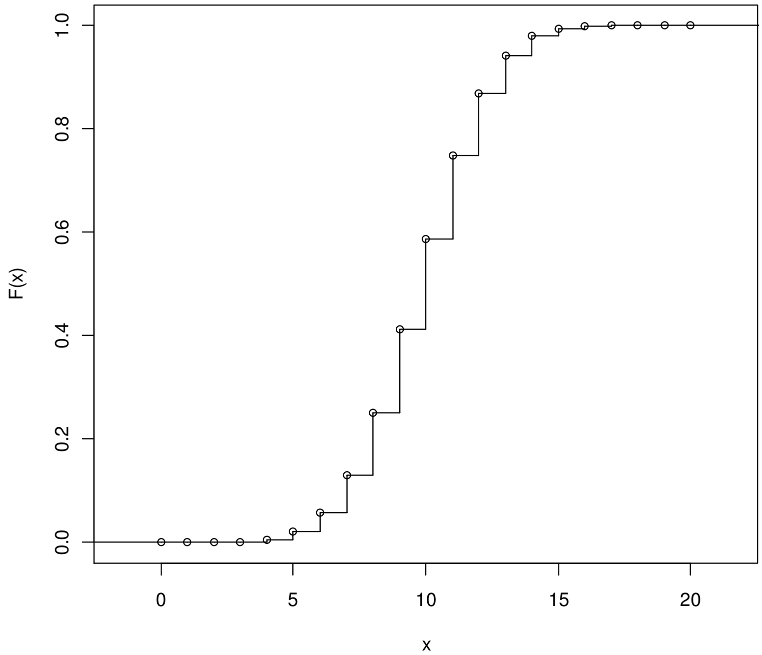


Figure 8 : Diagramme en escaliers d'une variable binomiale
Source : Pr BARANKANIRA Emmanuel

L'appellation de cette loi vient du fait que $C_n^k p^k (1-p)^{n-k} = C_n^k p^k q^{n-k}$ est le terme général due binôme de Newton de la forme :

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k} \tag{3.18}$$

En remplaçant a par p et b par q dans cette formule, il vient :

$$(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+1-p)^n = 1^n = 1 \tag{3.19}$$

L'espérance mathématique et la variance d'une variable aléatoire X qui suit une loi binomiale de paramètres n et p valent respectivement (**à démontrer**) [17] :

$$\begin{aligned} E(X) &= np \\ Var(X) &= np(1-p) \end{aligned} \tag{3.20}$$

Ces moments peuvent aussi se déduire de ceux d'une variable de Bernoulli de paramètre *p* dans la mesure où la loi binomiale est une somme de lois indépendantes de Bernoulli :

$$\begin{aligned} E(X) &= np \\ Var(X) &= np(1-p) \end{aligned}$$

L'écart-type de X vaut alors :

$$\sigma_X = \sqrt{np(1-p)} \tag{3.21}$$

Pour pouvoir utiliser cette loi, les conditions suivantes doivent être satisfaites :

- La variable étudiée doit être binaire ;
- Les tentatives doivent être indépendantes les unes des autres (tirage avec remise) ;
- La probabilité de succès doit rester constante après chaque épreuve ;
- L'échantillon doit être au sort [EAS] ;
- La taille n de l'échantillon doit être négligeable par rapport à la taille N de la population de laquelle il a été extrait $\left(\frac{n}{N} < 10\ \%\right)$.

Lorsque cette dernière condition n'est pas respectée, il faut utiliser une autre loi appelée loi hypergéométrique (voir plus loin). Remarquons que si deux variables aléatoires indépendantes X_1 et X_2 suivent des lois binomiales de paramètres respectifs (n_1, p) et (n_2, p), alors la variable somme $X_1 + X_2$ suit elle aussi une loi de binomiale de paramètres $n_1 + n_2$ et p**. C'est l'additivité ou la convolution des lois.**

Exemple :

Une expérience aléatoire consiste lancer 2 dés. Soit X une variable représentant la somme des points amenés par les 2 dés. L'expérience est répétée à 4 reprises. Le succès consiste en l'obtention d'une somme égale à 7.

a) Calculer les probabilités de tous les événements élémentaires
b) Vérifier que la somme de ces probabilités vaut 1
c) En déduire *E(X)* et *Var(X)*

Solution :

+	1	2	3	4	5	6
1						7
2					7	
3				7		
4			7			
5		7				
6	7					

Considérons les événements :

Succès : « Obtenir une somme égale à 7 »

Échec : « Obtenir une somme différente de 7 »

Les probabilités de ces événements sont :

$$p(succès) = p = \frac{6}{36} = \frac{1}{6}$$

$$p(Echec) = q = 1 - p = \frac{5}{6}$$

Les essais sont indépendants. Le nombre d'essais vaut $n = 4$. Il s'agit de la loi binomiale. Soit X la variable aléatoire qui représente le nombre de succès au cours de ces 4 essais. Cette suit une loi binomiale de paramètres $n = 4$ et $p = \frac{1}{6}$:

$$X \sim Bin\left(4, \frac{1}{6}\right)$$

La probabilité que la variable aléatoire X prenne une valeur k est donnée par :

$$p(X = k) = C_4^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{4-k} \quad avec\ k \in \{0,1,2,3,4\}$$

a) Les probabilités des événements élémentaires sont :

$$p(X = 0) = C_4^0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^4 = \left(\frac{5}{6}\right)^4 = 0,482 \cong 0,48$$

$$p(X = 1) = C_4^1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^3 = 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 = 4\left(\frac{125}{1296}\right) = \frac{500}{1296} = 0,85 \cong 0,9$$

$$p(X = 2) = C_4^2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = 6\left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = 6\left(\frac{1}{36}\right)\left(\frac{25}{36}\right) = \frac{25}{216} = 0,115 \cong 0,12$$

$$p(X = 3) = C_4^3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right) = 4\left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right) = \frac{20}{1296} = 0,0159 \cong 0,01$$

$$p(X = 4) = C_4^4 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^0 = \frac{1}{1296} = 0,00077 \cong 0,00$$

b) La somme des probabilités se calcule en cherchant les probabilités individuelles.

$$\sum_{i=0}^{4} p(X = x_i) = p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) + p(X = 4)$$

$$= 0,48 + 0,39 + 0,01 + 0,12 + 0,00 = 1$$

c) Les moments de X sont :

$$E(X) = np = 4\left(\frac{1}{6}\right) = \frac{2}{3} = 0,666 \cong 0,67$$

$$Var(X) = np(1-p) = 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right) = \frac{20}{36} = 0,555 \cong 0,56$$

Exercices d'application :

1) Dans une entreprise de fabrication d'ordinateurs, la probabilité de trouver un ordinateur qui ne fonctionne pas connecte set de 20 %. On décide d'examiner 5 ordinateurs tirés au hasard dans cette population d'ordinateurs. Quelle est la probabilité :

a) que les 5 ordinateurs ne fonctionnent pas correctement ?
b) que les 5 ordinateurs fonctionnent correctement ?
c) qu'au moins 1 de 5 ordinateurs ne fonctionnent pas correctement ?
d) que 3 ordinateurs au plus ne fonctionnent pas correctement ?

2) Des chambres à air sont produites en série et 5 % ont des défauts. Un garagiste en achète 10.
   a) Quelle est la probabilité que les 10 soient en bon état ?
   b) On suppose qu'il annule sa commande si plus de 2 articles ont des défauts. Quelle est la probabilité qu'il annule sa commande ?

3.5. Loi hypergéométrique

Considérons une urne contenant N boules dont N_A boules de type A dans laquelle le tirage de boules se fait sans remise (tirages successifs sans remise, donc indépendants). Dans ce cas, il u aura un moment où l'urne sera vide (tirage exhaustif). Comme il y a N_A boules de type A, alors il y aura $N - N_A$ boules de type B (non A). Soit $X(\omega)$ la variable aléatoire représentant le nombre de boules de type A tirés à l'issue de l'événement élémentaire ω. Le tirage des boules de type A et B se fait avec les probabilités $\boldsymbol{p}$ et $\boldsymbol{q}$ respectivement [17] :

$$\begin{aligned} p &= \frac{N_A}{N} \\ q &= 1 - p = \frac{N - N_A}{N} \end{aligned} \tag{3.22}$$

Le succès consiste donc en l'obtention d'une boule blanche. Un échantillon de ***n*** boules est prélevé de l'urne parmi lesquelles n_A boules de type A boules de type A et donc $n-n_A$ boules de type B (**Figure 9**). Contrairement à la loi binomiale où le tirage se fait avec remise, la variable aléatoire *X* qui représente le nombre de boules blanches qu'il est possible de trouver parmi *n* boules tirées sans remises suit une loi hypergéométrique de paramètres *N*, *n* et *p*. Elle porte le nom de variable hypergéométrique et se note [17] :

$$X \sim H\left(N, n, p\right) \tag{3.23}$$

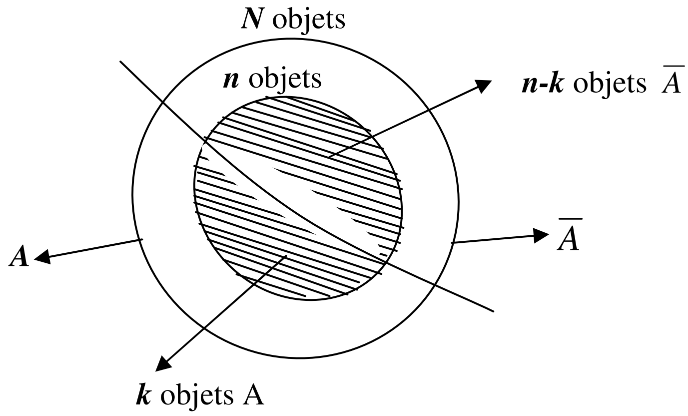


Figure 9 : Schéma du tirage exhaustif
Source : Pr BARANKANIRA Emmanuel

La probabilité qu'une variable aléatoire X qui suit une loi hypergéométrique de paramètres *N*, *n* et *p* s'écrit [8,9,11] :

$$p\left(X=k\right)=\frac{\begin{pmatrix} Np \\ k \end{pmatrix}\begin{pmatrix} Nq \\ n-k \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}=\frac{C_{Np}^{k}C_{Nq}^{n-k}}{C_{N}^{n}}, k\in\left\{0,1,2,...,n\right\} \tag{3.24}$$

Dans cette formule, il faut s'assurer que Np et Nq sont des nombres naturels.

L'espérance mathématique et la variance d'une variable aléatoire X qui suit une loi hypergéométrique de paramètres *N*, *n* et *p* sont respectivement (**à démontrer**) :

$$\begin{aligned} &E\left(X\right)=np \\ &Var\left(X\right)=np\left(1-p\right)\frac{N-n}{N-1} \end{aligned} \tag{3.25}$$

L'écart-type de *X* vaut alors :

$$\sigma_X=\sqrt{p\left(1-p\right)\frac{N-n}{N-1}} \tag{3.26}$$

L'espérance mathématique d'une loi hypergéométrique est la même que celle d'une loi binomiale. Cependant, la variance d'une loi binomiale est multipliée par le facteur d'exhaustivité appelé aussi coefficient d'exhaustivité qui fait que la composition de l'urne s'épuise puisque le tirage se fait sans remise :

$$e = \frac{N-n}{N-1} \tag{3.27}$$

La loi hypergéométrique converge vers la loi binomiale lorsque *N* est suffisamment grand devant *n* ou lorsque *N* tend vers l'infini :

$$H(N; n; p) \xrightarrow{N \to +\infty} Bin(n, p)$$

En effet :

$$\lim_{N \to +\infty} Var(X) = \lim_{N \to +\infty} \left[np(1-p)\frac{N-n}{N-1} \right] = np(1-p)$$

Si N>>>n, alors :

$$Var(X) = np(1-p)\frac{N-n}{N-1} \simeq np(1-p)$$

Exemple :

Dans une classe de 25 étudiants ,10 sont des tricheurs. Parmi eux, 5 étudiants sont choisis au hasard (sans remise).

a) Quel est le nombre de tricheurs espérons- nous trouver parmi eux ?
b) Quelle est la probabilité qu'il y ait moins de non tricheurs que de tricheurs ?

Solution :

N=25
n=5

Soient les événements :

Succès : « Être tricheur » $\Rightarrow p(succes) = p = \frac{10}{25} = \frac{2}{5}$

Échec : « Être non tricheur » $\Rightarrow p(Ehec) = q = 1 - p = 1 - \frac{2}{5} = \frac{3}{5}$

$$NP = 25\left(\frac{2}{5}\right) = 10$$

$$Nq = 25\left(\frac{3}{5}\right) = 15$$

Soit X la variable aléatoire qui désigne le nombre des tricheurs dans l'échantillon de 5 étudiants. Cette variable suit une loi hypergéométrique de paramètres N=25, n=5 et $p=\frac{2}{5}$:

$$X \sim H\left(25;5;\frac{2}{5}\right)$$

La probabilité la variable aléatoire X prenne une valeur donnée k vaut :

$$P(X=k)=\frac{C_{10}^{k}C_{15}^{k}}{C_{25}^{5}} \text{ avec } k\in\{0,1,2,3,4,5\}$$

a) Le nombre espéré de tricheurs vaut :

$$E(X)=np=5\left(\frac{2}{5}\right)=2$$

Cela veut dire que le nombre espéré de tricheurs est de 2 parmi les 5 étudiants.

La probabilité d'y trouver moins de non tricheurs que de tricheurs vaut :

$$\begin{aligned}p(X\geq 3)&=p(X=3)+p(X=4)+p(X=5)\\&=\frac{C_{10}^{3}C_{15}^{2}}{C_{25}^{5}}+\frac{C_{10}^{4}C_{15}^{1}}{C_{25}^{5}}+\frac{C_{10}^{5}C_{15}^{0}}{C_{25}^{5}}\\&=0{,}2371542+0{,}05928854+0{,}00474083\\&=0{,}3011858\\&=0{,}30\end{aligned}$$

3.6. Loi de Poisson

Appelée loi des petites probabilités, la loi de Poisson régit les événements rares. Autrement dit, elle permet de décrire les évènements dont les chances de réalisation sont faibles. Il s'agit d'une variable de comptage. Si, par exemple, un enquêteur s'assoit sur un pont pendant un intervalle de temps de donné (10h-11h par exemple) pour compter le nombre de voitures qui passent sur ce pont, alors la variable aléatoire qui représente ce nombre de personnes est une variable qui suit une loi de Poisson (variable de Poisson) dont le paramètre ($\lambda>0$ est le nombre moyen d'occurrences) est égal au nombre moyen de personnes qui passent sur ce pont sur n journées observées. Elle prend une infinité de valeurs $X(\Omega)=\mathbb{N}=\{0,1,2,...\}$. Cette variable décrit le nombre de fois qu'un événement rare se produit par unité de temps.

Autres exemples :

- Nombre de défauts sur un tissu
- Nombre de retours de marchandise au service après-vente
- Nombre de crédits avec un taux d'intérêt annuel inférieur à 10 %
- Nombre de bonus chez Lumitel au cours d'un mois

Elle se note [17] :

$$X \sim P(\lambda) \tag{3.28}$$

Exemples de variable de Poisson :

- Nombre d'appels dans une cabine téléphonique dans un intervalle de temps court ;
- Nombre mensuel d'accidents sur une route dans un intervalle de temps court ;
- Nombre annuel de naissances dans une petite municipalité ;

La probabilité qu'une variable aléatoire qui suit une loi de Poisson de paramètre λ prenne une valeur donnée k s'écrit [11,17] :

$$p(X=k)=\frac{\lambda^k}{k!}e^{-\lambda} avec\ k=0,1,\cdots \text{ ie } k\in \mathbb{N} \tag{3.29}$$

Ses moments (espérance mathématique et variance) sont (**à démontrer**) :

$$\begin{aligned} E(X)&=\lambda \\ Var(X)&=\lambda \end{aligned} \tag{3.30}$$

C'est l'unique loi pour laquelle l'espérance mathématique et la variance sont égales. Cette loi est dite **équidispersée**.

Exemple :

Dans une étoffe de tissu, on rencontre en moyenne 4 défauts sur 120m. Le tissu est débité en coupons de 5m. Combien de tissus sans défauts espérons-nous trouver dans un stock de 100 coupons.

Solution :

Soit X la variable aléatoire qui désigne le nombre de défauts sur un coupon. Cette variable suit une loi de Poisson de paramètre égal à :

$$\lambda=\frac{4(5)}{120}=\frac{20}{120}=\frac{1}{6}$$

La probabilité d'observer k défauts sur un coupon vaut :

$$p(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}\ avec\ k\in \mathbb{N}$$

soit

$$p(X=k)=\frac{\left(\frac{1}{6}\right)^k}{k!}e^{-\frac{1}{6}}$$

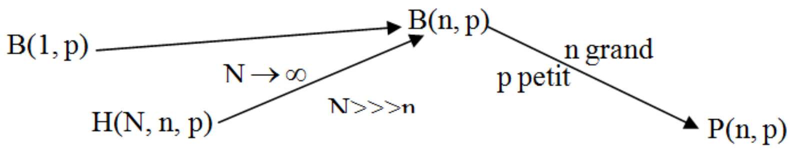
La probabilité d'observer 0 défaut sur un coupon (un coupon sans défaut) vaut :

$$p(X=0)=\frac{\left(\frac{1}{6}\right)^{0}}{0!}e^{-\frac{1}{6}}=e^{-\frac{1}{6}}=0,8464\cong 0,85$$

Dans un stock de 100 coupons, on espère y trouver 85 coupons sans défauts.

Convergence de la loi binomiale vers la loi de Poisson :

$$\lim_{n\to+\infty} B(k;n;p)=\frac{\lambda^k}{k!}e^{-\lambda}=p(k;\lambda)\ avec\ \lambda=np$$



Source : Pr BARANKANIRA Emmanuel

Une loi binomiale *Bin(n, p)* converge vers une loi de Poisson *P(λ)* si *p* est petit (en pratique inférieur à 0,10) et *n* assez grand (supérieur à 50). Dans ce cas, le paramètre de la loi de Poisson est *λ=np*. La loi de Poisson facilite les calculs par rapport à la loi binomiale. De plus, pour cette dernière, les valeurs prises par la variable aléatoire ne peuvent, en aucun cas, dépasser la valeur de *n*, alors que l'approximation par la loi de Poisson autorise des valeurs supérieures.

Le tableau 11 montre la comparaison des valeurs des probabilités pour une loi binomiale de paramètres *n* et *p* et une loi de Poisson de paramètre $\lambda=np$. Lorsque *n* augmente et *p* devient petit de façon que $\lambda=np$ soit de quelques unités (1,2,3,4), alors utiliser une loi binomiale ou une loi de Poisson revient au même.

Tableau 11 : Convergence d'une loi binomiale vers une loi de Poisson

k	$Bin\left(k,20,\frac{1}{10}\right)$	$Bin\left(k,1000,\frac{2}{1000}\right)$	$P(k,\lambda)$
0	0,122	0,133	0,135
1	0,272	0,271	0,275
2	0,285	0,273	0,271
3	0,190	0,180	0,180
4	0,089	0,090	0,090
5	0,132	0,035	0,036
6	0,009	0,011	0,012
7	0,002	0,003	0,003
⋮	⋮	⋮	⋮

En remplaçant p par $\frac{\lambda}{n}$ dans l'expression de la probabilité pour une loi binomiale, ce qui veut dire $\lambda = np$, il vient :

$$\begin{aligned}
C_n^k p^k (1-p)^{n-k} &= C_n^k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!}\left[\frac{n!}{(n-k)!}\left(\frac{1}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}\right] \\
&= \frac{\lambda^k}{k!}\left[\frac{n!}{(n-k)!}\left(\frac{1}{n}\right)^n (n-\lambda)^{n-k}\right]
\end{aligned}$$

En utilisant la formule de Stirling :

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{3.31}$$

il vient :

$$\frac{n!}{(n-k)!} \approx \sqrt{\frac{n}{n-k}}\,\frac{n^n}{(n-k)^{n-k}}\,e^{-k}$$

Le passage à la limite donne :

$$\lim_{n\to+\infty}\frac{n!}{(n-k)!}\left(\frac{1}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \approx \lim_{n\to+\infty}\sqrt{\frac{n}{n-k}}\left(\frac{n-\lambda}{n-k}\right)^{n-k} e^{-k} = \lim_{n\to+\infty}\left(\frac{n-\lambda}{n-k}\right)^{n-k} e^{-k}$$

Or,

$$\left(\frac{n-\lambda}{n-k}\right)^{n-k} = e^{(n-k)ln\frac{n-\lambda}{n-k}} = e^{(n-k)ln\left(\frac{n-\lambda}{n-k}-1+1\right)}$$

Le développement donne :

$$ln\left(\frac{n-\lambda}{n-k}-1+1\right) = \frac{n-\lambda}{n-k}-1+O\left[\left(\frac{n-\lambda}{n-k}-1\right)^2\right] = \frac{-\lambda+k}{n-k}+O\left[\left(\frac{-\lambda+k}{n-k}\right)^2\right]$$

Donc :

$$e^{(n-k)ln\frac{n-\lambda}{n-k}} = e^{-\lambda+k+(n-k)O\left[\left(\frac{-\lambda+k}{n-k}\right)^2\right]}$$

et

$$\lim_{n\to+\infty} e^{(n-k)\ln\frac{n-\lambda}{n-k}} e^{-k} = \lim_{n\to+\infty} e^{-\lambda+k+(n-k)O\left[\left(\frac{-\lambda+k}{n-k}\right)^2\right]} e^{-k} = e^{-\lambda+k}e^{-k} = e^{-\lambda}$$

La figure 10 montre les distributions des lois binomiale $(n=10;\ p=0,5)$ et de Poisson $(\lambda = np = 5)$. Les courbes ne sont pas proches.

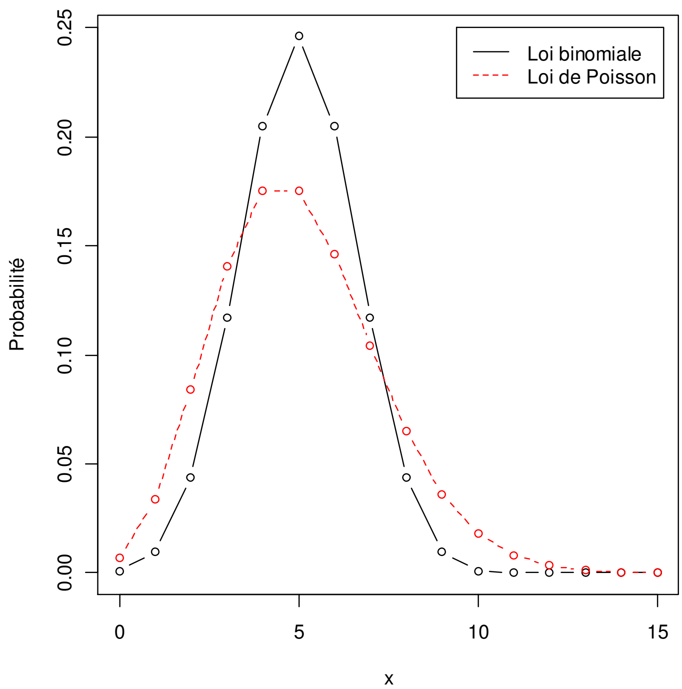


Figure 10 : Superposition des lois binomiale et de Poisson
Source : Pr BARANKANIRA Emmanuel

Remarquons que si deux variables aléatoires indépendantes X_1 et X_2 suivent des lois de Poisson de paramètres respectifs λ_1 et λ_2, alors la variable $X_1 + X_2$ suit elle aussi distribuée une loi de Poisson de paramètre $\lambda_1 + \lambda_2$. **C'est la convolution des lois.**

3.7. Loi géométrique

Les épreuves successives indépendantes sont réalsiées jusqu'à ce qu'un événement d'intérêt se produise. Pour chaque épreuve, il y a succès (avec une probabilité égale à *p*) ou échec (avec une probabilité égale à *q=1-p*). La variable aléatoire qui désigne le nombre d'épreuves réalisées jusqu'à la réalisation de cet événement particulier (avant l'obtention d'un succès) est une variable qui suit une loi géométrique ou de Pascal de paramètre *p* [17] :

$$X \sim G(p)$$

Elle est appelée variable géométrique ou variable de Pascal. Si cet événement est réalisé au k$^{\text{ème}}$ tirage, cela veut dire qu'il y a eu échec au cours des *(k-1)* premiers essais quitte à écrire la probabilité qu'une variable aléatoire *X* prenne une valeur donnée *k* comme suit :

$$p(X=k)=p^k(1-p)^{k-1} \text{ avec } k\in\mathbb{N}^* \tag{3.32}$$

L'espérance mathématique d'une variable aléatoire X qui suit une loi géométrique de paramètre p sont respectivement données par (**à démontrer**) :

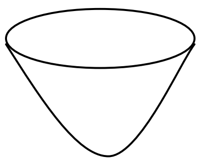
$$\begin{aligned} E(X)&=\frac{1}{p}\\ Var(X)&=\frac{1}{p^2}\end{aligned} \tag{3.33}$$

L'écart-type de *X* vaut alors :

$$\sigma_X=\frac{\sqrt{1-p}}{p} \tag{3.34}$$

Exemple :

Considérons une urne contenant 10 boules dont 4 blanches et 6 noires. Le succès consiste en l'obtention d'une boule blanche.



Considérons les événements :

Succès : « Obtention d'une boule blanche » $\rightarrow p=\frac{4}{10}$

Échec: « Obtention d'une boule noire » $\rightarrow q=1-p=\frac{6}{10}$

Soit *X* la variable aléatoire qui désigne le nombre d'épreuves réalisées avant l'obtention d'une boule blanche.

La probabilité d'obtenir une boule blanche au 5$^{\text{ème}}$ tirage est vaut :

$$p(X=5)=\frac{4}{10}\left(\frac{6}{10}\right)^4=0,4\times(0,6)^4\simeq 0,05$$

3.8. Loi binomiale négative

Comme la loi binomiale est une généralisation de la loi de Bernoulli, la loi binomiale négative ou loi de Polyà est une généralisation de la loi géométrique. En effet, répétons des épreuves successives indépendantes jusqu'à ce qu'un certain nombre *n* d'événements *A* soient réalisés. Pour la loi binomiale, le nombre d'épreuves est **fixé (connu)** et le nombre de réalisations est **aléatoire**. Cependant, pour la loi binomiale négative, le nombre de réalisations est fixé et le nombre d'épreuves nécessaires pour les obtenir est aléatoire. Pour la loi binomiale, la variable d'intérêt est le nombre de succès alors que pour la loi binomiale négative, la variable représente le nombre d'échecs. Considérons la variable aléatoire Y représentant le nombre d'épreuves effectuées. L'événement $Y = y$ est alors représentée par une suite d'événements $\underbrace{A\bar{A}AA...\bar{A}A\bar{A}}_{y-1}A$ comportant $(n-1)$ réalisations de l'événement *A* au cours des $(y-1)$ premières épreuves se terminant par l'événement *A*. Autrement dit, au lieu de calculer la probabilité de voir apparaitre k fois l'événement de probabilité p au cours de n épreuves identiques et indépendants, la loi binomiale négative permet de déterminer la probabilité d'être obligé d'effectuer n expériences identiques et indépendantes pour obtenir r fois l'événement de probabilité p. La variable Y suit une loi binomiale négative de paramètres *n* et *p*. Elle est appelée variable binomiale négative. La loi géométrique ou de Pascal est la loi binomiale négative avec n=1.

La probabilité qu'une variable aléatoire Y qui suit une loi binomiale négative de paramètres *n* et *p* prenne une valeur *y* donnée s'écrit [17] :

$$p(Y = y) = C_{y-1}^{n-1} p^{y-1} (1-p)^{y-n} \text{ avec } y \geq n \tag{3.35}$$

Les moments d'une variable aléatoire Y qui suit une loi binomiale négative de paramètres *n* et *p* se déduisent de ceux d'une variable *X* qui suit une loi géométrique ou de Pascal de paramètres *n* et *p* comme suit (**à démontrer**) :

$$\begin{aligned} E(Y) &= nE(X) = \frac{n}{p} \\ Var(Y) &= nVar(X) = \frac{nq}{p^2} = \frac{n(1-p)}{p^2} \end{aligned} \tag{3.36}$$

L'écart-type de *X* vaut alors :

$$\sigma_X = \frac{\sqrt{n(1-p)}}{p} \tag{3.37}$$

Illustration :

Une pièce de monnaie dont la probabilité d'obtenir pile est égale à *p* est lancée. Supposons que les lancers soient indépendants les uns des autres. Soit $r \geq 1$. Notons par *X* la variable aléatoire égale au nombre des faces obtenues avant le r$^{\text{ème}}$ pile. Quelle est la loi de *X* ? La variable *X* prend ses valeurs dans $\mathbb{N}$. Fixons $k \in \mathbb{N}$. Alors, si $X = k$ (r+k lancers sont faits amenant *r* piles et *k* amenant faces. Calculer $p(X = k)$ revient à calculer la probabilité que, dans la répétition de $(r+k)$ épreuves de

Bernoulli avec la probabilité p de succès, r fois le succès (piles) et k fois l'échec (face) sont obtenues. Fixons Ω un tel événement élémentaire, c'est-à-dire une suite de $(r+k)$ lancers comprenant r piles, k faces et se terminant par un pile. Par indépendance des lancers, la probabilité d'un tel événement élémentaire est $p^r(1-p)^k$.

Il reste donc à compter combien il y a de tels lancers. Puisque le dernier lancer est un pile, il suffit de compter le nombre de lancers amenant k faces parmi $(r+k-1)$ lancers (ou dans un arbre de Bernoulli $(r+k-1)$ épreuves. Combien de chemins comportent exactement k échecs) ? Par définition des coefficients binomiaux, il y en a exactement $\binom{r+k-1}{k} = C_{r+k-1}^{k}$.

Il vient que :

$$p(X=k) = \binom{r+k-1}{k} p^r (1-p)^k \tag{3.38}$$

Sous le schéma de Bernoulli (*épreuves* identiques et indépendants), le souhait est d'obtenir n succès. Considérons la variable aléatoire discrète X qui représente le nombre d'épreuves indépendantes k nécessaires à l'obtention de n succès. X suit une loi binomiale négative de paramètre n et p avec :

$$p(X=k) = C_{k=1}^{n-k} p^n q^{k-n} avec\ k, n \in \mathbb{N}, k \geq n \tag{3.39}$$

L'espérance associée à une loi binomiale négative est [17] :

$$E(X) = \frac{n}{p} \tag{3.40}$$

La variance associée à une loi binomiale négative est :

$$Var(X) = n\frac{q}{p^2} = n\frac{(1-p)}{p^2} \tag{3.41}$$

Pour renforcer les idées, dans le cas de la loi binomiale négative, le nombre de succès n est connu et ce qui est cherché est le nombre d'épreuves k nécessaires pour obtenir les n succès. Ainsi, le dernier événement est connu car les épreuves cessent avec l'obtention du nième succès et $(n-1)$ objets sont choisis parmi $(k-1)$.

Exemple :

Pour étudier le domaine vital d'une population de poissons, des émetteurs radio sont fixés au niveau de la nageoire dorsale. Après une légère anesthésie locale suite à divers aléas, considérons que 30 % des poissons équipés ne sont pas repérés par la suite. En considérant qu'au maximum 15 poissons doivent être suivis pour avoir des résultats statistiques acceptables, le variable aléatoire X qui désigne le nombre de poissons qui doivent être équipés suit une loi binomiale.

Cela se note par $X \to BN(15;0{,}70)$. En posant comme hypothèse que les causes des pertes de liaison radio soient suffisamment nombreuses pour assurer l'indépendance entre chaque épreuve, la probabilité d'être obligé d'équiper 20 poissons est :

$$P(X=20)=\frac{19!}{14!5!}(0{,}70)^{15}(0{,}30)^{5}=0{,}13$$

La figure 11 montre la superposition des lois binomiale de paramètres n=10 et p=0,5, de Poisson de paramètre $\lambda=5$ et binomiale négative de paramètres n=5 et p=0,5.

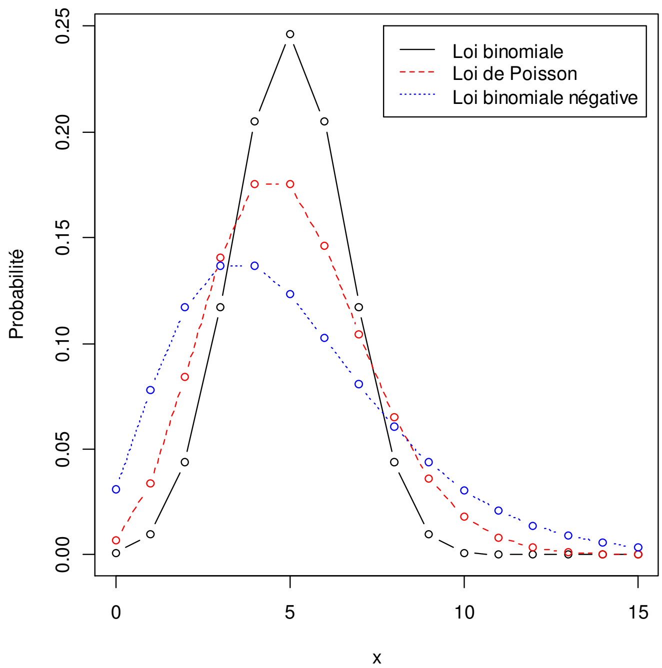


Figure 11 : Superposition des lois binomiale, de Poisson et binomiale négative
Source : Pr BARANKANIRA Emmanuel

3.9. Exercices d'application

Exercice 1 :

Il a été constaté que le nombre de bateaux qui mouillent dans un port est de 90 bateaux par mois.

Calculez la probabilité que :

a) aucun bateau ne mouille pendant 1 jour.
b) le nombre de bateaux qui mouillent est au moins égal à 3 pendant un jour.

Exercice 2 :

On admet que le nombre de défauts X sur le verre d'une ampoule de télévision suit une loi de Poisson de paramètre $\lambda = 4$. Calculez la probabilité des événements suivants :

a) Il n'y aucun défaut sur l'ampoule.
b) Il y a plus de 2 défauts sur l'ampoule.
c) Le nombre de défauts est compris entre 3 et 7 (bornes incluses)

Exercice 3 :

Sachant que le nombre de moyen de communications téléphoniques reçus par un standard entre 10h et 11h est de 1,8/min, calculez la probabilité qu'entre 10h53min et 10h54min, il y ait aucun appel, 1 appel, 2 appels, au moins 2 appels, plus de 2 appels, 2, 3 ou 4 appels ?:

Chapitre 4. Lois usuelles continues

4.1. Loi uniforme continue

Considérons deux réels a et b qui sont des bornes d'un intervalle. Une variable aléatoire réelle X suit la loi uniforme sur l'intervalle $[a,b]$ si sa densité de probabilité est de la forme [17] :

$$f(x)=\begin{cases} k \text{ } si \text{ } x\in[a,b] \\ 0 \text{ } sinon \end{cases} \tag{4.1}$$

Comme $f(x)$ est une densité de probabilité, elle doit vérifier la relation :

$$\begin{aligned}
&\int_{-\infty}^{+\infty} f(x)\,dx=1 \\
&\Leftrightarrow k\int_a^b dx=1 \\
&\Leftrightarrow k[c]_a^b=1 \\
&\Leftrightarrow k(b-a)=1 \\
&\Leftrightarrow k=\frac{1}{b-a}
\end{aligned}$$

Autrement dit, $f(x)$ est constante sur $[a,b]$ et nulle ailleurs :

$$\begin{aligned} f(x)&=\begin{cases} \dfrac{1}{b-a} \text{ } si \text{ } x\in[a,b] \\ 0 \text{ } sinon \end{cases} \\ &=I_{[a,b]}(x) \end{aligned} \tag{4.2}$$

où $I_{[a,b]}(x)$ est la variable indicatrice qui vaut 1 si $x\in[a,b]$ et 0 sinon. D'où, la densité de probabilité a pour valeur sur l'intervalle $[a,b]$ l'inverse de sa longueur.

Si $x>a$, la fonction de répartition de X est :

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\int_{-\infty}^{x} 0dt=[cte]_{-\infty}^{x}=0$$

Si $a\le x\le b$, la fonction de répartition de X est :

$$F(x)=0+\int_a^x f(t)\,dt=\frac{1}{b-a}\int_a^x dt=\frac{1}{b-a}[t]_a^x=\frac{x-a}{b-a}$$

Si $x < b$, la fonction de répartition de *X* est :

$$F(x) = 1 + \int_b^x f(t)\,dt = 1 + \int_b^x 0dt = 1 + [cte]_b^x = 1$$

La fonction de répartition de *X* est alors :

$$F(x) = \begin{cases} 0 \; si \; x < a \\ \dfrac{x-a}{b-a} \; si \; a \le x \le b \\ 1 \; si \; x > b \end{cases} \tag{4.3}$$

L'espérance mathématique et la variance d'une variable aléatoire *X* qui suit une loi uniforme continue sur l'intervalle $[a,b]$ sont respectivement données par (**à démontrer**) :

$$\begin{aligned} E(X) &= \frac{b+a}{2} \\ Var(X) &= \frac{(b-a)^2}{12} \end{aligned} \tag{4.4}$$

L'écart-type de *X* vaut alors :

$$\sigma_X = \frac{b-a}{2\sqrt{3}} \tag{4.5}$$

4.2. Loi exponentielle

La loi exponentielle permet d'étudier, par exemples, le temps d'attente d'un client dans une banque ou dans un hôpital, le temps au mariage, le temps d'attente jusqu'au nouveau travail, le temps de survie d'un malade, le temps de survie des ampoules électriques, la durée d'hospitalisation et le temps que les marchandises passent au stock.

La variable aléatoire *X* d'intérêt représente le temps d'attente entre deux événements successifs ou encore une durée de vie. Une variable aléatoire positive *X* suit une loi exponentielle de paramètre θ positif si elle admet la loi de densité définie par [16] :

$$f(x) = \begin{cases} \theta e^{-\theta x} si \; x \ge o \\ 0 \; ailleurs \end{cases} = \theta e^{-\theta x} I_{[0,+\infty]}(x) \tag{4.6}$$

Une variable aléatoire *X* qui suit une loi exponentielle de paramètre θ s'écrit [17] :

$$X \sim \varepsilon(\theta) \tag{4.7}$$

Si $x < 0$, la fonction de répartition de *X* est :

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^x 0dx = [cte]_{-\infty}^x = 0$$

Si $x \geq 0$, la fonction de répartition de X est :

$$F(x) = 0 + \int_0^x f(t)dt = \theta \int_0^x e^{-\theta t} dt = -\left[e^{-\theta t}\right]_0^x = -e^{-\theta x} + 1 = 1 - e^{-\theta x}$$

La fonction de répartition de X est alors :

$$F(x) = \begin{cases} 0 \ si \ x < 0 \\ 1 - e^{-\theta x} \ si \ \mathrm{x} \geq 0 \end{cases} \tag{4.8}$$

La figure 12 ci-après montre la fonction de répartition d'une variable exponentielle.

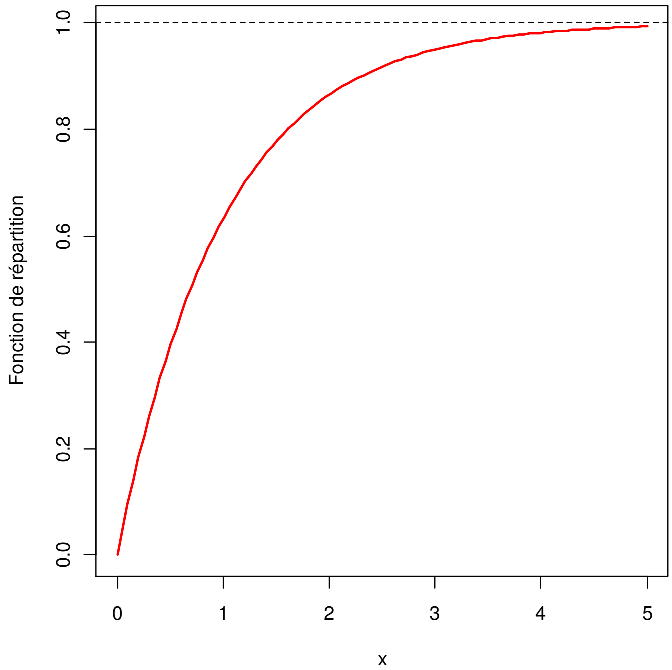


Figure 12 : Fonction de répartition d'une variable exponentielle
Source : Pr BARANKANIRA Emmanuel

L'espérance mathématique et la variance d'une variable aléatoire X qui suit une loi exponentielle de paramètre θ sont respectivement données par (**à démontrer**) [17] :

$$E(X) = \frac{1}{\theta}$$
$$Var(X) = \frac{1}{\theta^2} \tag{4.9}$$

L'écart-type de X vaut alors :

$$\sigma_X = \frac{1}{\theta} \tag{4.10}$$

4.3. Loi normale ou loi de Gauss-Laplace

Appelée loi de Gauss-Laplace, la loi normale est la loi de probabilité la plus importante. En effet, toutes les lois convergent vers la loi normale. Cette loi décrit les fluctuations aléatoires autour de l'espérance mathématique ou moyenne espérée. Elle est appelée couramment loi de Gauss ou loi de Laplace.

Une variable aléatoire continue X suit une loi normale de paramètres μ (moyenne) et $\sigma > 0$ (écart-type positif) si elle admet pour densité la fonction densité de probabilité f définie sur $\mathbb{R}$ par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (4.11)$$

Cette variable se note [17] :

$$X \sim N(\mu, \sigma) \qquad (4.12)$$

Comme $f(x)$ est une densité de probabilité, elle vérifie :

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

$$\Leftrightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

$$\Leftrightarrow \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \sigma\sqrt{2\pi}$$

Cette dernière relation est une intégrale de Poisson de type I.

La figure 13 montre la courbe densité de probabilité pour une variable aléatoire X qui suit une loi normale de paramètres $\mu = 3$ et $\sigma = 1,5$.

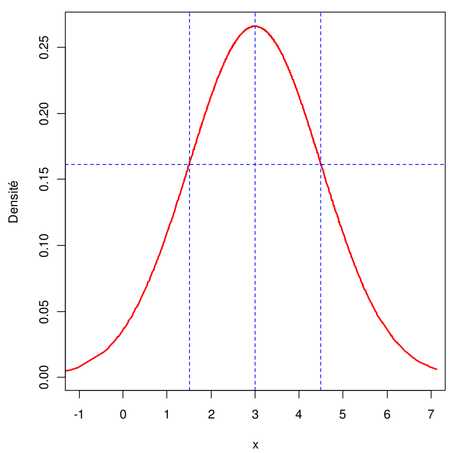


Figure 13 : Loi normale de Gauss-Laplace
Source : Pr BARANKANIRA Emmanuel

Le constat est que la courbe représentative de f :

- est une courbe dite en cloche ;
- est symétrique par rapport à la droite d'équation $x = \mu$;
- atteint son maximum en $\frac{1}{\sigma\sqrt{2\pi}}$;
- a le mode, la médiane et la moyenne égaux ;
- a l'aire totale sous la courbe égale à $\int_{-\infty}^{+\infty} f(x)dx = 1$.

L'espérance mathématique et la variance d'une variable aléatoire *X* qui suit une loi normale de paramètres μ et σ sont respectivement données par (**à démontrer**) [17] :

$$\begin{aligned} E(X) &= \mu \\ Var(X) &= \sigma^2 \end{aligned} \tag{4.13}$$

L'écart-type de *X* vaut :

$$\sigma_X = \sigma \tag{4.14}$$

4.4. Loi normale centrée–réduite ou loi normale standard

En faisant le changement de variable :

$$U = \frac{X - \mu}{\sigma} \; avec \; X \sim N(\mu, \sigma) \tag{4.15}$$

son espérance mathématique est :

$$E(U) = E\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma} E(X - \mu) = \frac{1}{\sigma}\left[E(X) - \mu\right] = \frac{1}{\sigma}(\mu - \mu) = \frac{1}{\sigma} \times 0 = 0 \tag{4.16}$$

et sa variance :

$$Var(U) = Var\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2} \operatorname{var}(X - \mu) = \frac{1}{\sigma^2} \operatorname{var}(X) = \frac{1}{\sigma^2} \times \sigma^2 = 1 \tag{4.17}$$

$$X \sim N(\mu, \sigma) \Rightarrow U = \frac{X - \mu}{\sigma} \sim N(0,1) \tag{4.18}$$

La densité de probabilité d'une variable aléatoire *U* qui suit une loi normale de moyenne 0 et d'écart-type 1, appelée variable centrée-réduite ou variable standardisée est :

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \tag{4.19}$$

L'intégration de cette densité donne :

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2} du = \sqrt{2\pi} \tag{4.20}$$

Le logarithme de la fonction $\varphi(u)$ est :

$$\ln \varphi(u) = -\ln\left(\sqrt{2\pi}\right) - \frac{1}{2}u^2$$

La dérivation du logarithme de la fonction $\varphi(u)$ est :

$$\frac{\varphi'(u)}{\varphi(u)} = -u$$
$$\Leftrightarrow \varphi'(u) = -u\varphi(u)$$
$$\Leftrightarrow \varphi'(u) = -\frac{u}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

En appliquant la transformation logarithmique à cette dérivée, il vient :

$$\ln \varphi'(u) = \ln\left(-\frac{u}{\sqrt{2\pi}}\right) - \frac{1}{2}u^2$$

La dérivation membre à membre donne :

$$\frac{\varphi''(u)}{\varphi'(u)} = \frac{1}{u} - u = \left(\frac{1-u^2}{u}\right)$$
$$\Leftrightarrow \varphi''(u) = -\left(\frac{u^2-1}{u}\right)\varphi'(u)$$
$$\Leftrightarrow \varphi''(u) = \frac{u^2-1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} = \left(u^2-1\right)\varphi(u)$$

La valeur maximale de $\varphi(u)$ est $\varphi(0) = \frac{1}{\sqrt{2\pi}} = 0,3989$ et les valeurs décroissent rapidement comme par exemples $\varphi(2) = 0,054$ et $\varphi(4) = 0,0001$. La fonction de répartition de la variable aléatoire *U* ne peut pas s'exprimer sous forme d'une fonction usuelle :

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}x^2} dx \tag{4.21}$$

En principe, chaque variable aléatoire a sa moyenne et son écart-type. Le rôle de la standardisation est d'éviter l'effet des unités de mesure. Cela conduit à une table unique de la loi normale. Les valeurs de $\Phi(u)$ permettent de calculer les probabilités pour une variable aléatoire de loi normale et sont données dans les tables statistisques. La fonction $\varphi(u)$ est paire :

$$\varphi(-u) = \varphi(u) \tag{4.22}$$

La distribution de $\Phi(u)$ est donc symétrique par rapport à une droite verticale qui passe par rapport au centre de distribution 0 (la moyenne d'une loi normale centrée et réduite est nulle), ce qui permet de calculer la probabilité :

$$p(U < -a) = p(U > a) = 1 - p(U \le a) \tag{4.23}$$

ou à l'aide de la fonction de répartition :

$$\Phi(-a) = 1 - \Phi(a) \tag{4.24}$$

4.4.1. Calcul des probabilités pour une loi normale

Si X est une variable aléatoire qui suit une loi continue, alors :

$$p(a < X < b) = p(X < b) - p(X < a) = \int_a^b f(x)\,dx \tag{4.21}$$

Puisque X obéit à une loi normale de moyenne μ et d'écart-type σ, alors la variable standardisée $U = \dfrac{X-\mu}{\sigma}$ suit une loi normale de moyenne 0 et d'écart-type 1. Une fonction de probabilité est toujours positive et l'aire de surface sous la courbe est toujours égale à 1. L'aire sous la courbe entre deux points donnés a et b représente la probabilité que la variable aléatoire X prenne des valeurs entre ces deux points [17] :

$$\begin{aligned} p(a < X < b) &= p\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) \\ &= p\left(\frac{a-\mu}{\sigma} < U < \frac{b-\mu}{\sigma}\right) \end{aligned} \tag{4.22}$$

La probabilité que la variable aléatoire X prenne une valeur donnée x_i est presque nulle :

$$p(X = x_i) \simeq 0 \tag{4.23}$$

La probabilité que la variable aléatoire X prenne une valeur inférieure (ou égale) à une valeur donnée a est donnée par la valeur de la fonction de répartition :

$$F(a) = p(X < a) = \int_{-\infty}^{a} f(x)\,dx \tag{4.24}$$

La standardisation d'une variable de loi normale permet d'utiliser une seule table statistique.

4.4.2. Comment utiliser la table statistique ?

Le calcul de la probabilité qu'une variable aléatoire de loi normale prenne une valeur inférieure à une valeur a donnée se fait à l'aide d'une table statistique à une décimale dans la colonne de gauche alors que la ligne de haut indique la seconde décimale (**Annexe 1**).

Exemple 1 : Soit X une variable aléatoire qui suit une loi normale de paramètres 0 et 1. Calculer :

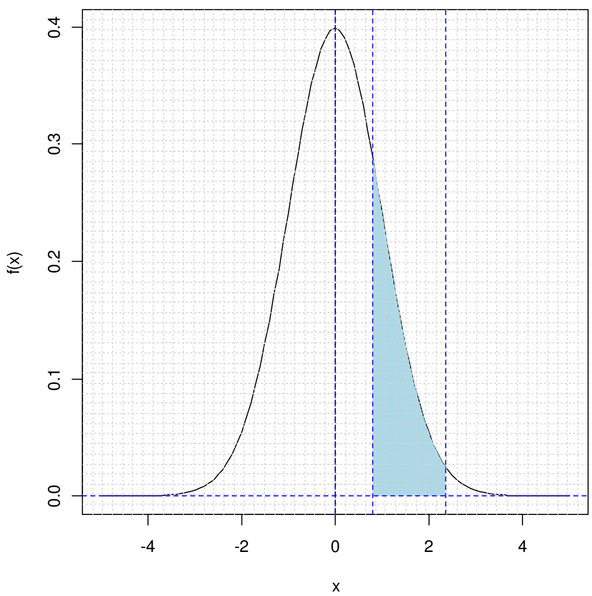
a) $p(0,80 < X < 2,35)$

b) $p(-2,11 < X < 0)$

Solution :

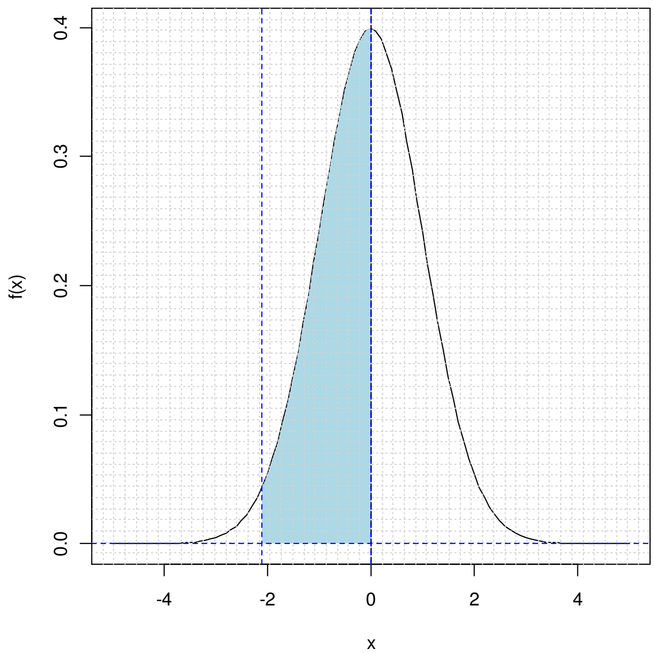
$$X \sim N(0,1)$$

a) $p(0,80 < X < 2,35) = ?$



$$\begin{aligned} p(0,80 < X < 2,35) &= p(X < 2,35) - p(X < 0,80) \\ &= 0,9906 - 0,7881 \\ &= 0,2025 \\ &\cong 0,20 \end{aligned}$$

b) $p(-2,11 < X < 0) = ?$



$$
\begin{aligned}
p(-2{,}11 < X < 0) &= p(X<0) - p(X<-2{,}11) \\
&= p(X<0) - p(X>2{,}11) \\
&= p(X<0) - \left[1 - P(X \le 2{,}11)\right] \\
&= p(X<0) - 1 + p(X \le 2{,}11) \\
&= 0{,}5000 - 1 + 0{,}9826 \\
&= 0{,}4826 \\
&\simeq 0{,}48
\end{aligned}
$$

Exemple 2 : Considérons une variable aléatoire X qui suit une loi normale de paramètres $\mu = 20$ et $\sigma = 10$: $X \sim N(20{,}10)$

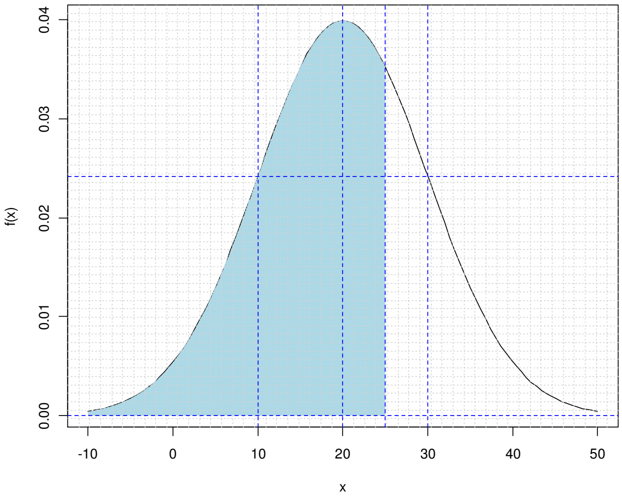
Calculer :

a) $p(X < 25)$

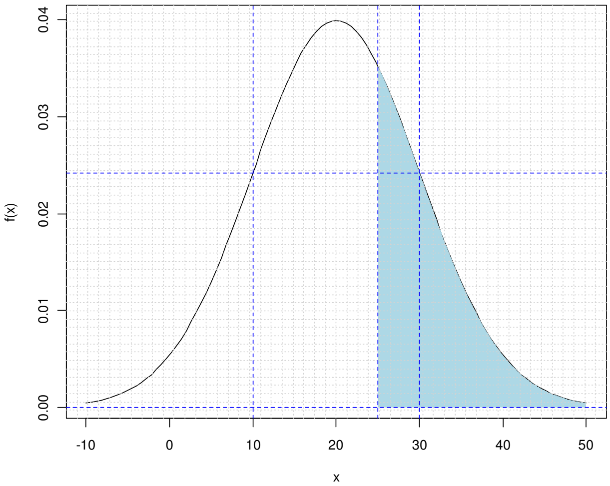
b) $p(X > 25)$

c) $p(15 < X < 18)$

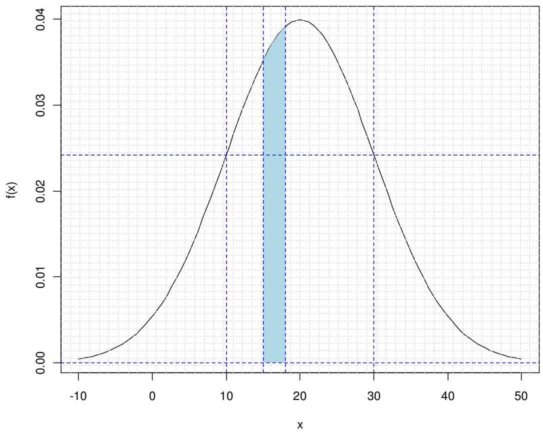
Solution :



$$
\begin{aligned}
p(X<25) &= p\left(\frac{X-20}{10}<\frac{25-20}{10}\right)\\
&= P(Z<0,5) \text{ avec } z=\frac{X-20}{10}\sim N(0,1)\\
&= 0,6915\\
&\cong 0,69
\end{aligned}
$$



$$
\begin{aligned}
p(X>25) &= 1-p\left(\frac{X-20}{10}\leq\frac{25-20}{10}\right)\\
&= 1-p(Z\leq 0,5)\\
&= 1-0,6915\\
&= 0,3085\\
&\cong 0,31
\end{aligned}
$$



$$
\begin{aligned}
p(15<X<18) &= p\left(\frac{15-20}{10}<\frac{X-20}{10}<\frac{18-20}{10}\right)\\
&= p(-0,5<Z<-0,2)\\
&= p(Z<-0,2)-p(Z<-0,5)\\
&= p(Z>0,2)-p(Z>0,5)\\
&= 1-p(Z\leq 0,2)-[1-p(Z\leq 0,5)\\
&= 1-p(Z\leq 0,2)-1+p(Z\leq 0,5)\\
&= p(Z\leq 0,5)-p(Z\leq 0,2)\\
&= 0,6915-0,5793\\
&= 0,1122\\
&\cong 0,11
\end{aligned}
$$

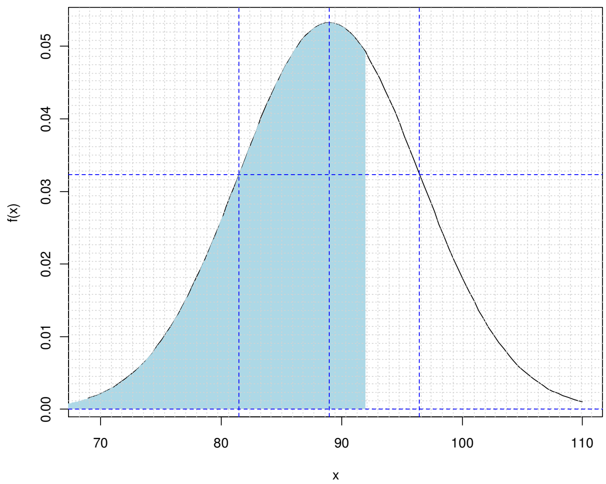
Exemple 3 : La demande d'un produit de consommation courante suit une loi normale de paramètre $\mu = 89$ et $\sigma = 7,5$. Calculer la probabilité que la demande soit :

a) inferieure à 92
b) supérieure à 98
c) comprise entre 92 et 98

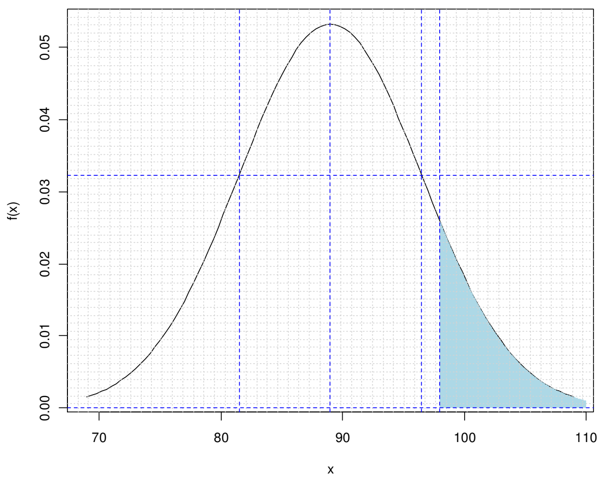
Solution

Soit X la variable aléatoire qui désigne la demande d'un produit de consommation courante :

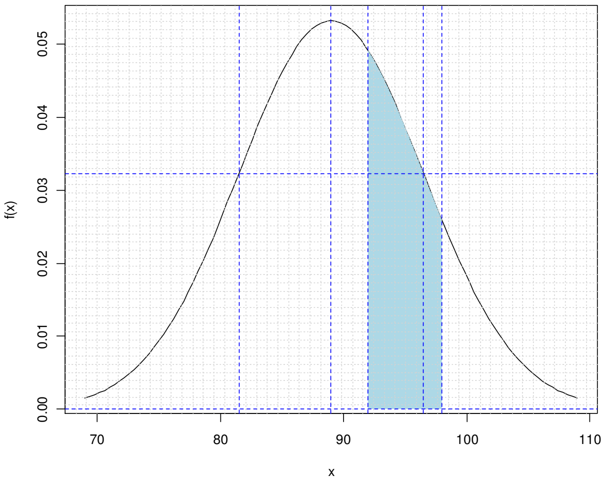
$$X \sim N(89;7,5)$$



$$\begin{aligned} p(X<92) &= p\left(\frac{X-89}{7,5}<\frac{92-89}{7,5}\right) \\ &= p(Z<0,4) \\ &= 0,6554 \\ &\cong 0,66 \end{aligned}$$



$$\begin{aligned} p(X>98) &= 1-p(X\le 98) \\ &= 1-p\left(\frac{X-89}{7,5}\le\frac{98-89}{7,5}\right) \\ &= 1-p(Z\le 1,2) \\ &= 1-0,8849 \\ &= 0,1151 \\ &\cong 0,12 \end{aligned}$$



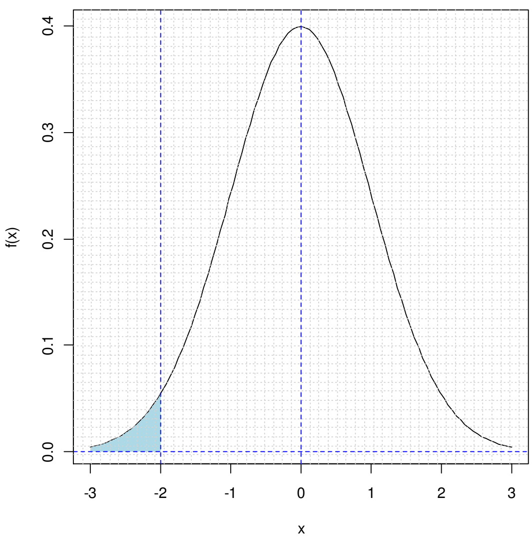
$$p(92 < X < 98) = p(X < 98) - p(X <$$
$$= p\left(\frac{X-89}{7,5} < \frac{98-89}{7,5}\right.$$
$$= p(Z < 0,12) - p(Z$$
$$= 0,8849 - 0,6554$$
$$= 0,2295$$
$$\cong 0,23$$

Exemple 4 : Soit U une variable de loi normale standard. Calculer :

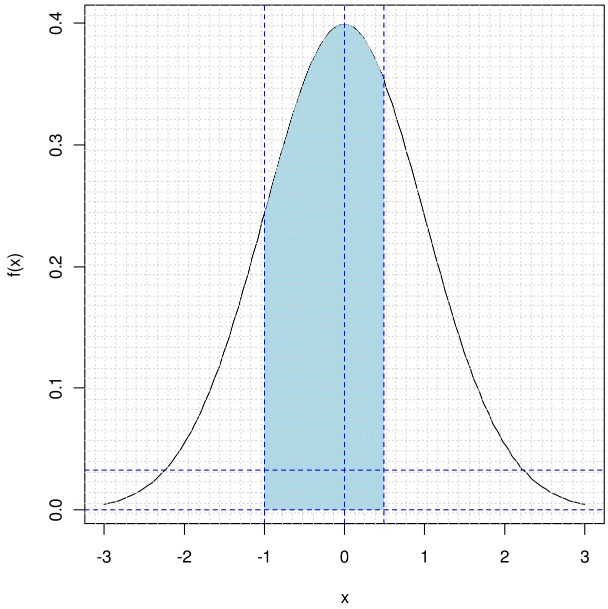
a) $p(U < -2)$ b) $p(-1 < U < 0,5)$ c) $p(4U \geq -3)$

Solution :

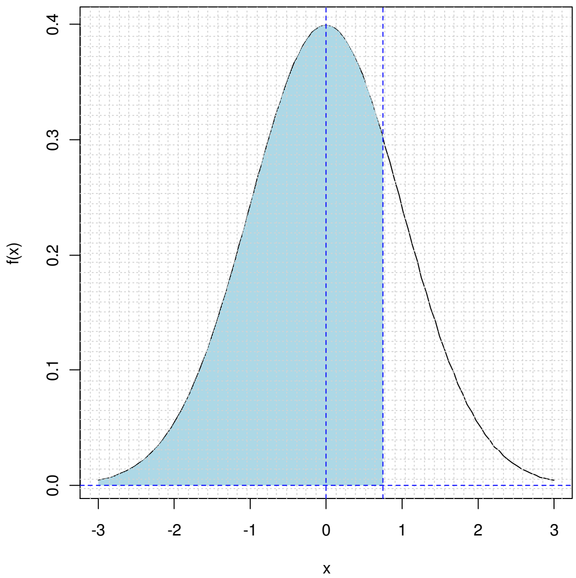
$$U \sim N(0,1)$$



$$p(U < -2) = p(U > 2)$$
$$= 1 - p(U \leq 2)$$
$$= 1 - 0,9772$$
$$= 0,0228$$
$$\cong 0,02$$



$$\begin{aligned} p(-1<U<0{,}5) &= p(U<0{,}5)-p(U<-1) \\ &= P(U<0{,}5)-P(U>1) \\ &= p(U<0{,}5)-\left[1-p(U\le 1)\right] \\ &= p(U<0{,}5)-1+p(U\le 1) \\ &= 0{,}6915-1+0{,}8413 \\ &= 0{,}5328 \\ &\cong 0{,}53 \end{aligned}$$



$$\begin{aligned} p(4U\ge -3) &= p(4U\le 3) \\ &= p\left(U\le \frac{3}{4}\right) \\ &= p(U\le 0{,}75) \\ &= 0{,}7734 \\ &\cong 0{,}77 \end{aligned}$$

Exercices d'application

Exercice 1 :

Déterminer, à l'aide de la table, les probabilités suivantes :

a) $p(-1,28<Z<1,28)$ b) $p(-1,64<Z<1,64)$ c) $p(-2<Z<2)$

d) $p(2,01<Z<1,4)$ e) $p(-2,01\le Z\le -1,4)$ f) $p(1,43\le Z\le 1,67)$

g) $p(1,365\le Z\le 2,36)$ h) $p(-2,01<Z<1,4)$ i) $p(1,215\le Z\le 3,213)$

NB : Pour les sous-questions g) et i), il suffit d'utiliser **l'interpolation linéaire**.

Exercice 2 :

Déterminez la valeur du quantile dans les cas suivants :

a) $p(0 < Z < z) = 0,2350$ b) $p(-z < Z < z) = 0,3631$ c) $p(z < Z < 1) = 0,2414$

d) $p(Z < z) = 0,8485$ e) $p(z \le Z \le 0) = 0,3444$ f) $p(z < Z \le 1) = 0,6314$

4.5. Distribution d'échantillonnage

Soient $X_1,...,X_n$ n variables aléatoires indépendantes et identiquement distribuées (iid), c'est-à-dire telles que $X_i \sim N(\mu,\sigma)$, ce qui conduit au fait que $E(X_i) = \mu, Var(X_i) = \sigma^2$. Considérons, en outre, la variable aléatoire qui est une somme de variables aléatoires $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ [18].

Dans ce cas (tirage avec remise), la moyenne de la distribution d'échantillonnage des moyennes est (**à démontrer**) [16,18] :

$$E(\bar{X}_n) = \mu$$
$$Var(\bar{X}_n) = \frac{\sigma^2}{n} \tag{4.25}$$

L'écart-type de la distribution d'échantillonnage des moyennes est donc :

$$\sigma_{\bar{X}_n} = \frac{\sigma}{\sqrt{n}} \tag{4.26}$$

Dans le cas du tirage sans remise, ces formules deviennent :

$$\begin{cases} E(\bar{X}_n) = \mu \\ Var(\bar{X}_n) = \dfrac{\sigma^2}{n}\dfrac{N-n}{N-1} \\ \sigma_{\bar{X}_n} = \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}} \end{cases} \tag{4.27}$$

4.6. Théorème central limite

Le théorème central limite (TCL) est aussi appelé théorème de la limite centrale. Soit (X_n)une suite de variables aléatoires indépendantes $(X_1, X_2, \ldots, X_n)$identiquement distribuées (iid) et de carré intégrable ou carré sommable, c'est-à-dire appartenant à l'ensemble des fonctions L^2.

Posons :

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \tag{4.28}$$

où $S_n = \sum_{i=1}^{n} X_i$. Alors la suite (Y_n) converge en loi vers une variable aléatoire de la loi normale standard, ce qui veut dire :

$$p(Y_n \leq t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}z^2} dz \tag{4.29}$$

Quel que soit la loi de probabilité, si un événement aléatoire est répété infiniment, alors sa moyenne finit par se comporter comme une loi normale. C'est ce théorème qui permet d'affirmer que la loi normale est la loi des phénomènes aléatoires.

Selon le théorème central limite (TCL) :

$$\overline{X}_n \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \text{ (dans le cas d'un tirage avec remise)} \tag{4.30}$$

$$\overline{X}_n \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right) \text{ (dans le cas d'un tirage sans remise)} \tag{4.31}$$

Dans ce cas de la distribution des proportions, considérons la variable aléatoire :

$$F_n = \frac{X}{n} \tag{4.32}$$

avec $X \sim Bin(n, p)$.

La moyenne en probabilité de la distribution d'échantillonnage des proportions est (**à démontrer**) :

$$E(F_n) = p$$

$$Var(F_n) = \frac{p(1-p)}{n} \tag{4.33}$$

L'écart-type de la distribution d'échantillonnage des moyennes est donc :

$$\sigma_{F_{nn}} = \sqrt{\frac{p(1-p)}{n}}$$

Dans le cas du tirage sans remise, ces formules deviennent :

$$\begin{cases} E(F_n) = p \\ Var(F_n) = \dfrac{p(1-p)}{n}\dfrac{N-n}{N-1} \\ \sigma_{F_{nn}} = \sqrt{\dfrac{p(1-p)}{n}}\sqrt{\dfrac{N-n}{N-1}} \end{cases} \quad (4.34)$$

Selon le théorème central limite (TCL) :

$$F_n \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right) \text{ (dans le cas avec remise)} \quad (4.35)$$

$$F_n \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}\right) \text{ (dans le cas sans remise)} \quad (4.36)$$

4.7. Loi log-normale

La loi log-normale, appelée aussi loi de Galton, est l'une des lois qui dérivent de la loi normale. Les autres lois de ce genre sont les loi de Student, loi du chi-deux (ou chi-carré) et loi de Fisher-Snedecor.

Il arrive souvent qu'une distribution ne soit pas normale. Dans ce cas, l'une des transformations à faire subir à la variable est la transformation logarithmique, les autres transformations étant par exemple arcsinus, inverse et racine carrée. Dans ce cas, si $\ln(X)$ est une variable normalement distribuée, alors la variable X positive suit une loi log-normale ou loi de Galton [17] :

$$X \sim LN(\mu, \sigma) \quad (4.37)$$

La fonction de répartition de la variable X est, pour x>0 :

$$\begin{aligned} F(x) &= p(X < x) \\ &= p(\ln X < \ln x) \\ &= p\left(\frac{\ln X - \mu}{\sigma} < \frac{\ln x - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{\ln x - \mu}{\sigma}\right) \end{aligned}$$

La fonction densité de probabilité de X pour x>0 est donc :

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} = \frac{1}{\sigma x}\varphi\left(\frac{\ln x - \mu}{\sigma}\right) \quad (4.38)$$

La loi log-normale porte aussi le nom de loi de l'effet proportionnel car l'additivité des effets conduit à la loi normale. C'est la multiplication des causes indépendantes ayant, chacune, un effet négligeable, qui conduit à la loi log-normale.

Par le même truchement que pour les lois précédentes, il est aisé de démontrer que l'espérance mathématique et la variance d'une variable aléatoire *X* qui suit une loi log-normale de paramètres μ et σ sont respectivement données par :

$$\begin{aligned} E(X) &= e^{\mu+\frac{\sigma^2}{2}} \\ Var(X) &= \left(e^{\sigma^2}-1\right)e^{2\mu+\sigma^2} \end{aligned} \tag{4.39}$$

4.8. Loi du khi-deux

La loi du χ^2 (khi-carré) permet de comparer des proportions (test d'homogénéité), de faire des tests de conformité d'une distribution observée à une distribution théorique, tester l'indépendance (test d'indépendance) de deux caractères qualitatifs ou la tendance des proportions.

Considérons $X_1,...,X_n$ n variables aléatoires indépendantes qui suivent une loi normale de moyenne 0 et d'écart-type. Ces variables aléatoires sont dites centrées-réduites. Une variable aléatoire définie par la somme des carrés des X_i par [17] :

$$X = X_1^2 + ... + X_n^2 = \sum_{i=1}^{n} X_i^2 \tag{4.40}$$

suit une loi du chi-deux de (Karl) Pearson à n degrés de liberté représentant le nombre de variables aléatoires indépendantes qui interviennent dans la définition de la variable :

$$X \sim \chi_n^{\ 2} \tag{4.41}$$

Pour de grands échantillons ($n \geq 30$), il est possible de montrer que $\sqrt{2X}-\sqrt{2n-1}$ a une distribution qui s'approche de la loi normale standard.

La loi du khi-deux à *n* degrés de liberté (ddl), notée χ_n^2 ou $\chi^2(n)$, est la loi $\gamma\left(\frac{n}{2},\frac{1}{2}\right)$ où *n* est un entier positif de densité pour x>0 [16] :

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{1}{2}} x^{\frac{n}{2}-1} \tag{4.42}$$

avec

$$\begin{aligned} &\Gamma(p) = \int_0^{+\infty} e^{-x}x^{p-1}dx = (p-1)\Gamma(p-1) \\ &\Leftrightarrow \Gamma(p) = (p-1)! \end{aligned}$$

Les moments de la loi Γ s'effectuent aisément par le changement de variables $y = \theta x$. Une variable aléatoire X suit une loi gamma de paramètres $p>0$ et $\theta > 0$ si c'est une variable aléatoire positive de densité de probabilité de la forme :

$$f(x) = \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1}, x \geq 0 \tag{4.43}$$

La fonction de répartition d'une variable aléatoire X qui suit une loi du chi-deux à n ddl ne s'écrit pas de manière explicite. Il existe une table statistique de la loi de Pearson qui permet de calculer une probabilité donnée en fonctions de quelques valeurs du degré de liberté (ddl) et de quelques seuils de probabilité (**Annexe 2**).

L'espérance mathématique et la variance d'une variable aléatoire X qui suit une loi gamma de paramètres $p>0$ et $\theta > 0$ sont (**à démontrer**) [17] :

$$\begin{aligned} E(X) &= \frac{p}{\theta} \\ Var(X) &= \frac{p(p+1)}{\theta^2} \end{aligned} \tag{4.44}$$

L'écart-type de X vaut alors :

$$\sigma_X = \frac{\sqrt{p(p+1)}}{\theta} \tag{4.45}$$

Pour une loi gamma avec $p = \frac{n}{2}$ et $\theta = \frac{1}{2}$ (loi du chi-deux avec n degrés de liberté), l'espérance mathématique, la variance et l'écart-type deviennent respectivement :

$$E(X) = \frac{P}{\theta} = \frac{\frac{n}{2}}{\frac{1}{2}} = \frac{n}{2} \times \frac{2}{1} = n$$

$$Var(X) = \frac{p(p+1)}{\theta^2} - \frac{p^2}{\theta^2} = \frac{p^2 + p - p^2}{\theta^2} = \frac{p}{\theta^2} = \frac{\frac{n}{2}}{\frac{1}{4}} = \frac{n}{2} \times \frac{4}{1} = 2n$$

$$\sigma_X = \sqrt{2n}$$

Il existe également un lien avec la loi normale qui explique son importance en statistique :

$$\text{Si } X \sim N(\mu, \sigma) \text{ alors } \frac{X-\mu}{\sigma} \sim N(0,1) \text{ et } \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi_1^2 \tag{4.46}$$

$$Z_i = \frac{X-\mu}{\sigma} \sim N(0,1) \Rightarrow Z_i \sim \chi_1^2 \Rightarrow T = \sum_{i=1}^{n} Z_i^2 \sim \chi_n^2 \qquad (4.47)$$

En d'autres mots, la somme $X = X_1 + ... + X_k$ de k variables aléatoires $X_1,...,X_k$ indépendantes distribuées selon une loi du khi-carré avec $n_1,...,\mathrm{n}_k$ degrés de liberté respectivement est une nouvelle variable aléatoire qui suit une loi du khi-carré avec $n = n_1 + ... + \mathrm{n}_k$ degrés de liberté. De même, si X_1, X_2 sont deux variables aléatoires indépendantes et si X_1 est distribuée selon une loi du khi-carré avec n_1 degrés de liberté et si $X = X_1 + X_2$ est distribuée selon une loi du khi-carré avec n degrés de liberté, avec $n > n_1$, alors X_2 est distribuée selon une loi du khi-carré avec $n - n_1$ degrés de liberté.

4.9. Loi de Student

Il a été signalé plus haut que la loi de Student est une loi dérivée de la loi normale. Cette loi est très utilisée dans les tests statistiques. Utilisée dans l'estimation des paramètres de population à partir des données observées sur l'échantillon, cette loi porte aussi le nom de loi Student-Fisher. Pour comprendre cette loi, considère une première variable aléatoire *X* qui est distribuée selon une loi normale centrée-réduite et une seconde variable *Y* qui lui est indépendante mais qui est distribuée selon une loi du χ^2à *n* degrés de liberté.

Dans ce cas, la variable définie par [17,19] :

$$Z = \frac{X}{\sqrt{\frac{Y}{n}}} = \sqrt{n}\frac{X}{\sqrt{Y}} \qquad (4.48)$$

suit une loi de Student à *n* degrés de liberté, notée t(*n*) :

$$Z \sim t(n) \qquad (4.49)$$

Le test de Student est un test utilisé lors des comparaisons des paramètres de population comme les moyennes (ou les proportions) et dans ce cas, le nombre de degrés de liberté vaut *n-1*. Sa généralisation peut être trouvée ailleurs [19].

Sachant que :

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \qquad (4.50)$$

il vient :

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \sim N(0,1) \qquad (4.51)$$

Cela implique aussi que, selon **le théorème central limite**, la moyenne d'une suite de variables aléatoires centrées et réduites et de carré intégrable, indépendantes et de même loi, une fois renormalisées par la racine carrée de *n*, converge en loi vers la loi normale standard [17] :

$$\sqrt{n}\,\frac{\overline{X}_n-\mu}{\sigma}\xrightarrow{loi}N(0,1) \qquad (4.52)$$

Dans le cas où le paramètre σ est inconnu, ce dernier est remplacé par l'écart-type empirique modifié, ce qui amène à considérer la variable :

$$\frac{S_n^2}{\sigma^2}=\frac{\frac{n-1}{\sigma^2}S_n^2}{n-1} \qquad (4.53)$$

dont le numérateur suit une loi $\chi^2(n-1)$. Le numérateur et le dénominateur (premier membre) sont des variables aléatoires et leur rapport est une nouvelle variable aléatoire qui suit une loi de Student à (n-1) ddl.

Soit *k* un entier strictement positif. Une variable aléatoire *X* suit une loi de Student à *n* degrés de liberté, notée $T(k)$, si *X* est une variable continue qui admet comme fonction densité de probabilité la fonction pour $x \geq 0$ [16] :

$$f(x)=\frac{1}{\sqrt{k\pi}}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}\frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}} \qquad (4.54)$$

La loi de Student a une fonction densité de probabilité qui est symétrique. De plus, si le degré de liberté *k* est grand (*k* dépassent 30), la loi de Student peut être approximée par la loi normale standard. Comme pour la loi du chi-carré, la fonction de répartition de la variable aléatoire *X* ne s'exprime pas de manière explicite. Il existe une table statistique de la fonction de répartition de *X* pour quelques valeurs de degré de liberté et quelques valeurs seuils (**Annexe 3**). Il est à remarquer que, pour *k=1*, la loi de Student devient la loi de *Cauchy*. Cette loi est, en fait, la loi de la variable aléatoire qui résulte du rapport entre deux variables aléatoires indépendantes qui suivent chacune la loi normale standard. De plus, par rapport à d'autres lois de probabilité, cette loi possède la particularité de n'admettre aucun moment.

Sa densité de probabilité est :

$$f(x)=\frac{1}{\pi\left(1+x^2\right)} \qquad (4.55)$$

L'espérance mathématique d'une variable aléatoire *X* qui suit une loi de Cauchy est donnée par :

$$E(X)=\int_{-\infty}^{+\infty} f(x)\,dx=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{xdx}{\left(1+x^2\right)}=\frac{1}{2\pi}\Big[\ln\left(1+x^2\right)\Big]_{-\infty}^{+\infty} \qquad (4.56)$$

qui est une intégrale divergente. L'espérance mathématique de *X* n'existe donc pas, ce qui en est de même pour la variance.

Les moments d'une variable aléatoire *X* qui suit une loi de Student à *k* degrés de liberté sont respectivement données par :

$$\begin{aligned} &E(X)=0\\ &Var(X)=\frac{k}{k-2}\ avec\ k>2\\ &\sigma_X=\sqrt{\frac{k}{k-2}}\ avec\ k>2 \end{aligned} \qquad (4.57)$$

Lorsque k tend vers l'infini, alors selon le théorème central limite, la loi de Student tend vers la loi normale standard [16] :

$$t(k)\xrightarrow{loi} N(0,1) \qquad (4.58)$$

4.10. Loi de Fisher-Snedecor

Les variances empiriques modifiée (corrigée) et non modifiée sont respectivement données par :

$$\begin{aligned} &S_n^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2\\ &S_n^{2'}=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2 \end{aligned} \qquad (4.59)$$

$$
\begin{aligned}
E\left(S_n^{2'}\right) &= E\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}E\left(X_i-\bar{X}_n\right)^2 \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\mu+\mu-\bar{X}_n\right)^2\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\mu\right)^2-2\left(X_i-\mu\right)\left(\bar{X}_n\right)+\left(\bar{X}_n-\mu\right)^2\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\mu\right)^2-2n\left(\bar{X}_n-\mu\right)^2+n\left(\bar{X}_n-\mu\right)^2\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\mu\right)-n\left(\bar{X}_n-\mu\right)^2\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\mu\right)^2\right]-E\left(\bar{X}_n-\mu\right)^2 \\
&= Var\left(X\right)-Var\left(\bar{X}_n\right) \\
&= \sigma^2-\frac{\sigma^2}{n} \\
&= \sigma^2\left(1-\frac{1}{n}\right) \\
&= \frac{n-1}{n}\times\sigma^2 \neq \sigma^2
\end{aligned}
$$

$S_n^{2'}$ est donc un estimateur biaisé de σ^2 .

Le biais de cet estimateur est :

$$
biais\left(S_n^{2'}\right)=E\left(\delta_n'^2\right)-\sigma^2=\frac{n-1-n}{n}\sigma^2=-\frac{1}{n}\sigma^2\neq 0 \tag{4.60}
$$

$$
\begin{aligned}
E\left(S_n^2\right) &= E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2\right] = \frac{1}{n-1}\sum_{i=1}^{n}E\left(X_i-\bar{X}_n\right)^2 \\
&= \frac{1}{n-1}E\sum_{i=1}^{n}\left(X_i-\mu+\mu-\bar{X}_n\right)^2 \\
&= \frac{1}{n-1}E\left\{\sum_{i=1}^{n}\left[\left(X_i-\mu\right)^2-2\left(X_i-\mu\right)\left(\bar{X}_n-\mu\right)+\left(\bar{X}_n-\mu\right)^2\right]\right\} \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left(X_i-\mu\right)^2-2n\left(\bar{X}_n-\mu\right)^2+n\left(\bar{X}_n-\mu\right)^2\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left(X_i-\mu\right)^2-n\left(\bar{X}_n-\mu\right)^2\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left(x_i-\mu\right)^2\right]-\frac{n}{n-1}E\left[\left(\bar{x}_n-\mu\right)^2\right] \\
&= \frac{n}{n-1}Var\left(x\right)-\frac{n}{n-1}Var\left(\bar{x}_n\right) \\
&= \frac{n}{n-1}\sigma^2-\frac{n}{n-1}\frac{\sigma^2}{n} \\
&= \frac{n\sigma^2}{n-1}-\frac{\sigma^2}{n-1} \\
&= \left(\frac{n}{n-1}-\frac{1}{n-1}\right)\sigma^2 \\
&= \frac{n-1}{n-1}\sigma^2 \\
&= \sigma^2
\end{aligned}
$$

S_n^2 n'est pas biaisé.

$$
Var\left(S_n^2\right)=\frac{\mu_4}{n}-\frac{n-3}{n\left(n-1\right)}\sigma^2=\sigma^2 \tag{4.61}
$$

où

$$
\mu_4=E\left[\left(x-\mu\right)^4\right]
$$

Pour la variance empirique, il vient :

$$
Var\left(S_n^{2'}\right)=\frac{\left(n-1\right)^2}{n^4}\left(\mu_4-\sigma^4\right)+2\frac{n-1}{n^3}\sigma^4
$$

Il est possible également de préciser le lien existant entre la moyenne et la variance empirique par le calcul de la covariance :

$$Cov\left(\bar{X}_n, S_n^{2'}\right) = \frac{(n-1)^2}{n^2}\mu^2 \quad (4.62)$$

et

$$Cov\left(\bar{X}_n, S_n^2\right) = \frac{\mu^3}{n} \quad (4.63)$$

Ainsi, pour une loi symétrique, il vient :

$$Cov\left(\bar{X}_n, S_n^2\right) = 0 \quad (4.64)$$

Considérons à présent deux échantillons $(X_1,...,X_n)$ et $(Y_1,...,Y_m)$ auxquels sont associées les variances empiriques $S_n^{2'}$ et $S_m^{2'}$. Le rapport $\frac{S_n^{2'}}{S_m^{2'}}$ a un numérateur et un dénominateur qui suivent, chacun, une loi du chi-carré. Il s'agit donc d'un rapport de deux variables aléatoires. Supposons que ces variables aient la même variance. Dans ce cas, considérons les variables définies par :

$$\begin{aligned} U &= \frac{(n-1)S_n^{2'}}{\sigma^2} \\ V &= \frac{(m-1)S_m^{2'}}{\sigma^2} \end{aligned} \quad (4.65)$$

Le rapport :

$$F = \frac{\frac{U}{n}}{\frac{V}{m}} \sim F(\mathrm{n}, \mathrm{m}) \quad (4.66)$$

est une nouvelle variable aléatoire appelée variable qui suit une loi de Fisher-Snedecor de paramètres *n* et *m*, c'est-à-dire à *n* degrés de liberté (au numérateur) et *m* degrés de liberté (au dénominateur).

Il est à remarquer que :

$$F' = \frac{\frac{V}{m}}{\frac{U}{n}} \sim F(m, n) \quad (4.67)$$

La loi de Fisher ou loi de Fisher-Snedecor, notée $F(n, m)$, est utilisée pour comparer deux variances observées et dans de nombreux tests d'analyse de la variance et de la covariance.

Contrairement à la loi de Student qui est symétrique en 0, la loi de Fisher-Snedecor est dissymétrique. Soit n et m deux entiers positifs. Une variable aléatoire X suit une loi de Fisher-Snedecor à n et m degrés de liberté si X est une variable continue et admet pour densité de probabilité la fonction f (x) suivante pour $x \geq 0$:

$$f(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}\left(\frac{n}{m}\right)^{\frac{n}{2}} \frac{x^{\frac{n-2}{2}}}{1+\left(\frac{n}{m}x\right)^{\frac{n+m}{2}}} \tag{4.68}$$

Comme pour la loi de Student, la fonction de répartition d'une variable aléatoire X qui suit une loi *F(n,m)* ne s'écrit pas de façon explicite. Il existe des tables statistiques de la fonction de répartition.

Les moments (espérance mathématique et variance d'une variable aléatoire X qui suit une loi de Fisher-Snedecor de paramètres m et n sont respectivement données par :

$$\begin{aligned} E(F) &= \frac{m}{m-2} \; avec \; m > 2 \\ Var(F) &= \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)} \; avec \; m > 4 \end{aligned} \tag{4.69}$$

Le mode d'une loi de Fisher $F(n,m)$ est unique et sa valeur est donnée par :

$$Mode = \frac{n-2}{n}\frac{m}{m+2} \; avec \; n > 2 \tag{4.70}$$

Les fractiles de cette loi sont tabulées pour certaines valeurs de n et m et pour certaines valeurs seuils (**Annexe 4**).

Chapitre 5 : Convergence en probabilité

5.1. Inégalité de Markov

Si X est une variable aléatoire réelle qui admet une espérance mathématique, alors l'inégalité de Markov stipule que $\forall \lambda > 0$:

$$p\left[X \geq \lambda E(X)\right] \leq \frac{1}{\lambda}$$
$$\Leftrightarrow p[X \geq \lambda] \leq \frac{E(X)}{\lambda} \tag{5.1}$$

Cette première forme de cette inégalité ne présente pas un intérêt pour $\lambda \leq 1$ puisque le majorant de la probabilité va dépasser 1, ce qui est en contradiction avec le premier axiome de Kolmogorov. La seconde forme de cette inégalité s'applique à $|X|^k$ pour tout k tel que $E\left(\left|X^k\right|\right)$ existe [17] :

$$p\left[\left|X^k\right| \geq \lambda\right] \leq \frac{E\left(\left|X^k\right|\right)}{\lambda} \tag{5.2}$$

En introduisant un nombre $\varepsilon > 0$ qui est tel que $\lambda = \varepsilon^k$, il vient :

$$p\left[|X| \geq \varepsilon\right] \leq \frac{E\left(\left|X^k\right|\right)}{\varepsilon^k} \tag{5.3}$$

5.2. Inégalité de Bienaymé-Tchebychev

L'inégalité de Bienaymé-Tchebychev est un cas particulier de celle de Markov. En effet, elle s'obtient en prenant *k=2* et en replaçant *X* par *X-E(X)* dans l'inégalité de Markov [17] :

$$p\left[\left|X - E(X)\right| \geq \varepsilon\right] \leq \frac{Var(X)}{\varepsilon^2} \tag{5.4}$$

Cette inégalité, qui n'a pas d'intérêt pour $\varepsilon \leq \sigma_X$, permet de relier la probabilité pour une variable aléatoire X de pouvoir s'écarter de son espérance mathématique à sa variance (paramètre de dispersion autour de la moyenne).

Pour une variable centrée-réduite, elle s'écrit $\forall a > 1$:

$$p\left[\left|\frac{X - E(X)}{\sigma_X}\right| \geq a\right] \leq \frac{1}{a^2} \tag{5.5}$$

5.3. Inégalité de Jensen

Considérons une fonction g réelle convexe définie sur un intervalle $I \subset \mathbb{R}$. Cette fonction est convexe si $g(\alpha x + (1-\alpha)) \leq \alpha g(x) + (1-\alpha) g(y), \forall x, y \in I$ avec $\alpha \in [0,1]$, $g^{"}(x) \geq 0$ telle que $E(X)$ et $E[g(X)]$ existent.

Dans ce cas, l'inégalité de Jensen s'écrit :

$$g[E(X)] \leq E[g(X)] \tag{5.6}$$

Cela veut dire que, au sens de *g*, l'ordonnée de la moyenne est plus petite que la moyenne des ordonnées.

5.4. Convergence presque sûre

La suite (X_n) converge presque sûrement vers la variable aléatoire X si il existe un événement N de mesure nulle tel que :

$$\lim_{n\to+\infty} X_n(\omega) = X(\omega), \forall \omega \in N^c \tag{5.7}$$

Cela se note :

$$X_n \xrightarrow{p.s} X \tag{5.8}$$

5.5. Convergence en probabilité

La suite (X_n) converge en probabilité vers X si $\forall \varepsilon > 0$ [17] :

$$\begin{aligned} &\lim_{n\to+\infty} p(|X_n - X| > \varepsilon) = 0 \\ \Leftrightarrow &\lim_{n\to+\infty} p(|X_n - X| < \varepsilon) = 1 \end{aligned} \tag{5.9}$$

Cela se note :

$$X_n \xrightarrow{p} X \tag{5.10}$$

Une autre façon de le dire est que, si $E(X_n) \to a$ et $Var(X_n) \to 0$ lorsque n tend vers l'infini, alors : $X_n \xrightarrow{p} a$.

Si les variables aléatoires (X_n) et X sont dans L^p, alors (X_n) converge dans L^p vers X si :

$$\lim_{n\to+\infty} E(|X_n - X|^p) = 0 \tag{5.11}$$

Dans ce cas, une suite (X_n) de variables aléatoires converge en moyenne d'ordre *p* si avec $0 < p < \infty$ vers la variable aléatoire *X*.

Cela se note :

$$X_n \xrightarrow{L^p} X \tag{5.12}$$

La convergence dans L^p et la convergence presque sûre entraînent la convergence en probabilité et la convergence en probabilité entraîne la convergence presque sûre.

5.6. Convergence en loi

Soient $F_1,...,F_n$ une suite de fonctions de répartition associées aux variables aléatoires réelles $X_1,...,X_n$ et *F* la fonction de répartition de la variable aléatoire réelle *X* qui est supposée continue. Les fonctions $F_n(x)$ et $F(x)$ sont respectivement définies comme suit :

$$\begin{aligned} F_n(x) &= p(X_n \leq x) \\ F(x) &= p(X \leq x) \end{aligned} \tag{5.13}$$

La suite (X_n) converge en loi ou en distribution vers *X* si :

$$\lim_{n\to+\infty} F_n(a) = F(a), \forall a \in \mathbb{R} \tag{5.14}$$

Cela se note :

$$X_n \xrightarrow{loi} X \text{ ou } X_n \xrightarrow{d} X \tag{5.15}$$

quand *n* tend vers l'infini, la variable aléatoire :

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{5.16}$$

converge en probabilité vers μ, c'est-à-dire :

$$\bar{X}_n \xrightarrow{p} \mu \tag{5.17}$$

5.7. Loi des grands nombres

Soient $X_1,\cdots,X_n$ n variables indépendantes et identiquement distribuées (iid) [17] :

$$X_i \sim N(\mu,\sigma) \tag{5.18}$$

Cela revient à considérer une suite $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles mutuellement indépendantes définies sur un espace probabilisé (Ω, A, p) et qui admettent des moments d'ordre 1 et 2 (espérance mathématique et variance) :

$$\begin{cases} E(X_i) = \mu \\ Var(X_i) = \sigma^2 \end{cases} \tag{5.19}$$

Considérons, en outre, la variable aléatoire :

$$S_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{5.20}$$

Ses moments sont :

$$E(\overline{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu \tag{5.21}$$

$$Var(\overline{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

La relation $E(\overline{X}_n) = \mu$ montre que $\overline{X}_n$ est un estimateur sans biais. En effet, soient $\hat{\theta}$ un estimateur de θ. $\hat{\theta}$ est un estimateur sans biais de $\hat{\theta}$ si $E(\hat{\theta}) = \theta$, ce qui veut dire que le biais est nul :

$$biais(\hat{\theta}) = E(\hat{\theta}) - \theta = 0 \tag{5.22}$$

L'inégalitéé de Bienaymé –Tchebychev s'écrit :

$$p\left[\left|\overline{X}_n - \mu\right| \geq \varepsilon\right] \leq \frac{Var(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n\to+\infty} 0 \tag{5.23}$$

Cela permet d'énoncer le théorème :

Théorème : loi faible des grands nombres

Si (X_n) est une suite de variables aléatoires mutuellement indépendantes qui admettent les mêmes moments d'ordre 1 et 2, c'est-à-dire avec tout entier, $E(X_n) = \mu$ et $Var(X_n) = \sigma^2$, alors quand $n \to \infty$:

$$\overline{X}_n \xrightarrow[P]{} \mu \tag{5.24}$$

C'est la convergence en probabilité. Ce théorème porte le nom de **loi faible des grands nombres** dans la mesure où les variables aléatoires ne doivent pas nécessairement être de même loi de probabilité.

Théorème : loi forte des grands nombres

Si (X_n) est une suite de variables aléatoires mutuellement indépendantes et de même loi de distribution qui admettent une espérance, alors lorsque n tend vers l'infini [17] :

$$\overline{X}_n \xrightarrow{p.s} \mu \tag{5.25}$$

Cela veut dire que la moyenne d'une suite de variables aléatoires converge presque sûrement vers la même limite que l'espérance mathématique de la moyenne. Ce théorème est aussi appelé loi des grands nombres. Le principe de la loi forte des grands nombres est donc donné sous certaines conditions (sur la dépendance, sur l'homogénéité et sur les moments). Ici, l'adjectif qualificatif fait référence à la nature de la convergence établie par ce théorème (convergence presque sûre) alors que la loi faible des grands nombres établie par Bernoulli fait référence uniquement à la convergence en probabilité.

5.8. Qualité des estimateurs

Par définition, un estimateur d'un paramètre θ est une application T_n qui, à un n-échanillon de variables aléatoires $(X_1,...,X_n)$ de la loi p_θ associe une variable aléatoire réelle (ou plusieurs dans le cas d'un paramètre multidimensionnel) dont il est possible de déterminer la loi de probabilité.

Autrement dit, un estimateur d'un paramètre de population est une variable aléatoire (définie par une expression mathématique) dont les valeurs fournissent les estimations ponctuelles du paramètre de population.

Ainsi donc, un estimateur $\hat{\theta}$ de θ est dit **sans biais** ssi :

$$E(\hat{\theta}) = \theta \tag{5.27}$$

Cela veut dire que son biais est nul :

$$B(\hat{\theta}) = 0 \Leftrightarrow E(\hat{\theta}) = \theta \tag{5.28}$$

Un estimateur est dit asymptotiquement sans biais si :

$$E_\theta(T_n) = \theta \xrightarrow{n\to+\infty} \theta \tag{5.29}$$

De plus, un estimateur $\hat{\theta}_1$ de θ est de **variance minimale** parmi tous les estimateurs possibles $\hat{\theta}$ de θ est celui qui a la plus petite variance :

$$Var(\hat{\theta}_1) \leq Var(\hat{\theta}_i) \tag{5.30}$$

Un estimateur $\hat{\theta}_n$ de θ est dit **convergent** ssi :

$$\hat{\theta}_n \xrightarrow[n\to+\infty]{p} \theta \tag{5.31}$$

Si T_n est un estimateur sans biais de θ, alors il vient l'inégalité :

$$P_\theta\left(\left|T_n-\theta\right|>\varepsilon\right)\to 0 \tag{5.32}$$

Un estimateur convergent est donc un estimateur sans biais et dont la variance est minimale si $n\to\infty$:

$$\left.\begin{array}{l} E(T_n)=\theta \\ Var(T_n)\xrightarrow{n\to+\infty}0 \end{array}\right\}\Rightarrow T_n\xrightarrow[p]{}\theta \tag{5.33}$$

Tout estimateur sans biais dont la variance $\to o$ est convergent.

Théorème 1

Ce résultat se déduit directement de l'inégalité Bienaymé –Tchebychev :

$$\begin{array}{l} p_\theta\left(\left|T_n-\theta\right|>\varepsilon\right)\le \dfrac{V_\theta(T_n)}{\varepsilon^2}\to 0,\forall\varepsilon>0,n\to\infty \\ V_\theta(T_n)=E_\theta\left[(T_n-\theta)^2\right]\to 0,\, n\to\infty \end{array} \tag{5.34}$$

Théorème2

Tout estimateur asymptotiquement sans biais dont la variance $\to 0$ est convergent :

$$\left.\begin{array}{l} E_\theta(T_n)=\theta \\ V_\theta(T_n)\to 0 \end{array}\right\}T_n\xrightarrow[p]{}\theta,n\to\infty \tag{5.35}$$

La propriété la plus désirable pour un estimateur est d'avoir **une faible** erreur quadratique moyenne :

$$\begin{aligned} EQM(\hat{\theta}) &= E\left[\left(\hat{\theta}-\theta\right)^2\right] \\ &= Var(\hat{\theta})+\left(E(\hat{\theta})-\theta\right) \\ &= Var(\hat{\theta})+B^2(\hat{\theta}) \end{aligned} \tag{5.36}$$

Parmi deux estimateurs $\hat{\theta}_1$ et $\hat{\theta}_2$ de θ, le plus efficace (le plus efficient) est celui qui a une petite variance : $\hat{\theta}_1$ est efficace que $\hat{\theta}_2$ ssi [17] :

$$Var(\hat{\theta}_1)\le Var(\hat{\theta}_2) \tag{5.37}$$

5.9. Vraisemblance d'un échantillon

Soient :

θ le vecteur de paramètres inconnus $\left(\theta \in \mathbb{R}^p\right)$

$f(x) = f(x;\theta)$ la loi de X

$(x_1, \cdots, x_n)$ une réalisation d'échantillon aléatoire

La vraisemblance de l'échantillon aléatoire s'écrit :

$$\ell(x_1, \cdots, x_n, \theta) = \prod_{i=1} f(x_i, \theta)$$

La lettre $\ell =$likelihood (en anglais) désigne la vraisemblance.

Le log –vraisemblance s'écrit :

$$L = \ln[\ell(x_i;\cdots;\theta)] = \sum_{i=1}^{n} \ln[f(x_i,\theta)] \qquad (5.38)$$

Si $f(x_i, \theta)$ est différentielle en θ, L est une fonction dérivable de θ et le score est sa dérivée :

$$\delta_n(\theta) = \frac{\partial L}{\partial \theta} = \frac{1}{\ln(x_1, \cdots, x_n, \theta)} \frac{\partial}{\partial \theta} \ell(x_1, \cdots, x_n, \theta) \qquad (5.39)$$

Le score est centré :

$$E(\delta_n(\theta)) = 0 \qquad (5.40)$$

La variance du score (si elle existe), appelée également information de FISHER apportée par l'échantillon sur θ vaut :

$$I_n(\theta) = E\left[\left(\delta_n(\theta)\right)^2\right] \qquad (5.41)$$

La quantité d'information de FISHER peut aussi s'écrire (si le domaine de définition de X ne dépend pas de θ et que cette quantité existe) :

$$I_n(\theta) = -E\left[\frac{\partial^2 L}{\partial \theta^2}\right] = E\left[\frac{\partial \delta_n(\theta)}{\partial \theta}\right] \qquad (5.42)$$

Soit aussi :

$$I_n(\theta) = nI_1\theta \qquad (5.43)$$

Soit I une statistique de l'échantillon et $g(t;\theta)$ sa loi, alors $I_n(\theta) \geq I_1(\theta)$ avec $I_T(\theta)$ l'information de FISHER apportée par T sur θ est [17] :

$$I_T(\theta) = -\left[\frac{\partial^2 \ln g(T;\theta)}{\partial \theta^2}\right] \quad (5.44)$$

La statistique T est dite **exhaustive** pour le paramètre θ si la distribution de l'échantillon aléatoire conditionnellement à $T = t$ ne dépend pas de θ.

Si T est exhaustive, alors :

$$I_n(\theta) = I_T(\theta) \quad (5.45)$$

Autrement dit, une statistique T de la loi g_θ est dite exhaustive pour le paramètre θ SSI il existe une fonction $k : \mathbb{R}^n \to \mathbb{R}$ telle que :

$$\ell_\theta(x_1, \cdots, x_n) = g_\theta(t) k(x_1, \cdots, x_n), \forall (x_1, \cdots, x_n) \in IR, \forall \theta \in H \quad (5.46)$$

5.10. Exercices d'application

Exercice 1 :

La proportion de tubes défectueux produits par une entreprise est de 2 %.

a) Quelle est la loi du nombre de tubes défectueux dans un échantillon de 200 tubes. Déterminer son espérance et sa variance.

b) Les conditions requises pour approcher cette loi par une loi de Poisson sont –elles satisfaites ?

Calculer la probabilité d'obtenir un nombre de tubes défectueux :

- nul
- égal à 5
- égal à 6
- supérieur ou égal à 10

Solution :

a) Il s'agit d'une loi binomiale de paramètres 200 et 0,02 (on répète 200 fois l'opération qui consiste à choisir au hasard un bulbe et on compte le nombre total de tubes défectueux).

Son espérance est :

$$E(X) = np = 4$$

et sa variance :

$$Var(X) = np(1-p) = 3{,}92$$

b) Le résultat « Choisir un tube défectueux est un événement rare », il serait donc logique d'approcher la loi par une loi de Poisson.

Comme $n \geq 30$, $p \leq 0{,}1$ et $np = 4 \leq 5$, les conditions énoncées dans le cours sont vérifiées et l'approximation de la loi binomiale par la loi de Poisson de paramètre peut être faite avec :

$$\lambda = np = 4$$

En utilisant l'approximation par une loi de Poisson de paramètre 4, il vient donc :

$$p(X = 0) = 0{,}0183 \simeq 0{,}02$$

$$p(X = 5) = 0{,}785 \simeq 0{,}79$$

$$p(X = 6) = 0{,}8893 \simeq 0{,}89$$

$$p(X = 10) = 0{,}997 \simeq 1$$

Exercice 2 :

Un ascenseur supporte une charge de 800kg. Le poids des utilisateurs est distribué selon la loi normale de paramètre $\mu = 80kg$ et $\sigma = 20kg$. Quel est nombre maximum de personnes que l'on peut autoriser à monter simultanément pour la probabilité de sa charge ne dépasse pas 0,001.

Solution :

Soit X la variable aléatoire qui représente le poids (masse) de n utilisateurs. X est la somme des n variables aléatoires indépendantes dont la loi de probabilité est une loi normale de paramètre $\mu = 80kg$ et $\sigma_2 = 400kg$. Par conséquent, la loi de X est aussi une loi normale de paramètres 80n et 400n.

On souhaite avoir :

$$p(X > 800) \leq 0{,}001$$

soit

$$X < 800 \geq 0{,}999$$

Mais

$$p(X < 800) = p\left(\frac{X - 80n}{\sqrt{400n}} < \frac{800 - 80n}{\sqrt{400n}}\right)$$

Donc, d'après la table de la loi normale centrée-réduite, on en déduit que :

$$\frac{800-80n}{\sqrt{400n}} \geq 3{,}08$$

soit

$$80n + 61{,}6\sqrt{n} - 800 \leq 0$$

soit enfin :

$$4n + 3\sqrt{n} - 40 \leq 0$$

Donc, en résolvant cette équation du second degré, il vient :

$$\begin{cases} \sqrt{n} \leq 2{,}8 \\ \sqrt{n} \leq 7{,}84 \end{cases}$$

Donc, le nombre maximal de personnes à admettre est 7.

Exercice 3 :

Indépendance des événements

Un entreprise réalise et commercialise des compositions florales ainsi que des produits pour le jardin. L'entreprise confectionne ses compositions florales avec des bulbes des fleurs qu'elle reçoit en grande quantité. Chaque bulbe peut présenter deux défauts que l'on désigne par défaut a et défaut b.

On prélève un bulbe au hasard dans un stock important. On note A l'événement : « Le bulbe présente par le défaut a » et on note B l'événement : « Le bulbe présente le défaut b ». On admet que les probabilités des événements A et B sont $p(A) = 0{,}015$ et $p(B) = 0{,}02$. On suppose que les deux événements A et B sont indépendants. On admettra que si A et B sont indépendants, alors $A, B, \overline{A}, \overline{B}$ sont indépendants deux à deux.

Calculer la probabilité des événements :

E_1 = « Le bulbe présente le défaut a et le défaut b » ;

E_2 = « le bulbe présente au moins un des deux défauts » ;

E_3 = « Le bulbe ne présente aucun des deux défauts ».

Solution :

Soient les événements :

A : « Le bulbe présent le défaut a »

B : « Le bulbe présente le défaut b »

E_1 : « Le bulbe présente le défaut a et le défaut »

E_2 : « Le bulbe présente au moins un des deux défauts »

E_3 : « Le bulbe ne présente aucun défaut »

$$p(A)\,p(B) = 0,015 \times 0,02 = 0,0003$$

$$p(E_1) = p(A \cap B) = p(A).p(B) = 0,015 \times 0,02 = 0,0003$$

Comme $p(A \cap B) = p(A).p(B)$ alors, les événements A et B sont indépendants.

$$\begin{aligned} P(E_2) = P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0,015 + 0,02 - 0,0003 \\ &= 0,0347 \\ &\simeq 0,03 \end{aligned}$$

E_3 est l'événement contraire de E_2

Donc

$$p(E_3) = 1 - p(E_2) = 1 - 0,0347 = 0,9653 \simeq 0,97$$

5.11. Exercices à faire

Exercice 1 :

Chercher l'estimateur du maximum de vraisemblance pour les familles de loi exponentielle de paramètre $\frac{1}{\theta}$.

Exercice 2 :

Chercher l'estimateur du maximum de vraisemblance des paramètres d'une loi normale.

Exercice 3 :

Sachant que la répartition des quotients intellectuels QI (rapport entre l'âge mental et l'âge réel) des circonscrits recrutés est une loi normale de moyenne 0,90 et d'écart-type 0,40. Calculez, pour un bataillon (1000personnes) :

a) Le nombre de circonscrits ayant un $QI < 1$
b) Le nombre de circonscrits ayant un $QI < 0{,}1$
c) Le nombre de circonscrits ayant un $QI > 1{,}4$
d) Le nombre de circonscrits ayant un QI compris entre 0,8 et 1,3

Exercice 4 :

Les ampoules de télévision produites par une firme A ont une durée de vie moyenne de 2500h avec un écart –type de 500h tandis que celles produites par la firme B ont une durée de vie moyenne de 2300h avec un écart – type de 800h. On prélève un échantillon de 300 ampoules chez A et de 200 chez B.

a) Calculez la probabilité pour que la durée de vie moyenne de l'échantillon de A ne soit supérieure de plus de 100h à la durée de vie moyenne de l'échantillon de B.
b) Calculez la probabilité pour que la durée de vie moyenne de l'échantillon de A soit supérieure de plus de 200h à la durée de vie moyenne de l'échantillon de B.

Chapitre 6. Inférence statistique, intervalle de confiance et tests d'hypothèses

Considérons le tableau 12 ci-après qui montre les paramètres de population à estimer et les statistiques calculées à partir de l'échantillon.

Tableau 12 : Paramètres et statistiques

	Paramètres	**Statistiques**	
Estimateur	Population	Echantillon	Dénomination
$\hat{\mu}$	μ	$\overline{x}$	Moyenne
$\hat{\pi}$	π	p	Proportion
$\hat{\sigma}^2$	σ^2	s^2	Variance
σ	σ	s	Déviation Standard
$\hat{N}$	N	n	Taille

Source : BARANKANIRA Emmanuel

6.1. Techniques d'échantillonnage probabilistes

Pour estimer les paramètres d'une population à partir d'un échantillon, il est très important de constituer l'échantillon dont les unités doivent être strictement tirées 3conditions au sort (hasard) :

- Établir la liste complète des unités constituant l'ensemble de la population à étudier, ce qui est parfois difficile à réaliser ;
- Effectuer le tirage des ces unités par un procédé assurant l'égale probabilité pour chaque unité statistique d'être choisie (échantillon aléatoire simple). Pour cela, un numéro est affecté à chaque unité et le tirage au sort de ces unités est effectué ;
- S'en tenir strictement aux unités désignées par le tirage au sort quelles que soient les difficultés pratiques.

6.1.1. Échantillonnage aléatoire simple

L'échantillonnage aléatoire simple (EAS) est le modèle d'échantillonnage le plus simple qui puisse être imaginé. Il consiste à considérer que, dans une population de taille N, tous les échantillons de *n* unités sont possibles avec la même probabilité. Cependant, l'échantillonnage aléatoire simple fournit un cadre de référence indispensable pour deux raisons : il sert en quelque sorte d'étalon et il constitue en général « la brique » élémentaire des plans usuels ; par exemple, l'échantillonnage stratifié et l'échantillonnage à deux degrés sont des assemblages de sondages simples.

Un échantillon est aléatoire si tous les individus de la population ont une probabilité non nulle et non certaine de faire partie de l'échantillon. Il est simple si les individus ont des chances égales de figurer dans l'échantillon. Le caractère aléatoire minimise le risque de non représentativité de l'échantillon. De plus, il est possible d'anticiper sur le degré de précision de l'échantillon obtenu, et ainsi éviter une enquête inutile. La méthode d'échantillonnage aléatoire simple permet la comparaison d'études similaires dans le temps. L'échantillonnage aléatoire simple peut se faire avec ou sans remise.

6.1.2. Échantillonnage systématique

L'échantillonnage systématique est une technique qui consiste à prélever des unités d'échantillonnage situées à intervalles égaux. Il consiste à tirer sur la base de sondage un individu sur ***k*** avec

$$k = \frac{\textit{Taille de la population}}{\textit{Taille de l'échantillon}} \tag{6.1}$$

Le choix du premier individu détermine la composition de tout échantillon. Si l'effectif total de la population $\boldsymbol{N}$ est connu, et que le souhait est de prélever un échantillon d'effectif n, l'intervalle entre deux unités successives à sélectionner est donné par :

$$k = \frac{N}{n} \tag{6.2}$$

Connaissant k, l'enquêteur choisit le plus souvent pour débuter un nombre aléatoire ***i*** compris entre ***1*** et ***k*** ; le rang des unités sélectionnées est alors : ***i, i+k, i+2k, i+3k, …***

L'échantillonnage systématique réduit le temps consacré à la localisation des unités sélectionnées**.** Si les éléments de la population se présentent dans un ordre aléatoire (pas de tendance), l'échantillonnage systématique est équivalent à l'échantillonnage aléatoire simple. Par contre, si les éléments de la population présentent une tendance, l'échantillonnage systématique est plus précis que l'échantillonnage aléatoire.

L'échantillonnage systématique peut être utilisé lorsque la taille de la population n'est pas connue. Dans cette situation, c'est le **pas** qui est établi à l'avance et par conséquent, la taille de l'échantillon sera aléatoire. Dans le cas où les variables étudiées présentent une dispersion forte dans la population, la précision des résultats d'un sondage aléatoire simple peut être améliorée par l'utilisation du **sondage aléatoire stratifié**.

6.1.3. Échantillonnage stratifié

L'échantillonnage stratifié est une technique qui consiste à subdiviser une population hétérogène, d'effectif N, en p sous-populations ou « strates » plus homogènes d'effectifs N_i de telle sorte que

$$N = N_1 + N_2 + \ldots + N_p \tag{6.3}$$

L'idée de base de la stratification est d'obtenir une estimation précise d'une moyenne de strate quelconque à partir d'un petit échantillon prélevé dans cette strate ainsi qu'une estimation précise pour l'ensemble de la population en combinant ces estimations.

Dans l'échantillonnage stratifié, la variance de l'estimateur ne comprend que la variation à l'intérieur des strates. Le degré de précision augmente avec le nombre de strate de la population car, plus elles sont nombreuses, plus les unités qu'elles contiennent sont nombreuses.

Pour un échantillon de taille ***n***, nous avons :

$$n = \sum_{h=1}^{H} n_h \tag{6.4}$$

Nous choisissons n_h individu dans la strate h de façon que le taux de sondage dans chaque strate soit le même que dans la population :

$$\frac{n_h}{N_h} = \frac{n}{N} \text{ pour toutes les strates } h \tag{6.5}$$

Pour réaliser un sondage aléatoire stratifié, il est nécessaire de disposer de la base de sondage. D'autre part, même si cette base de sondage est disponible, le coût de l'enquête peut être prohibitif en raison des coûts de déplacement lorsque la population est dispersée géographiquement et que l'enquête est réalisée par enquêteur à domicile. Dans ce cas, l'enquêteur fait recours à deux autres méthodes de sondage aléatoire: le **sondage en grappes** et **le sondage à plusieurs degrés**.

6.1.4. Échantillonnage à plusieurs degrés

Les sondages à plusieurs degrés utilisent une succession de regroupements des unités statistiques pour tirer l'échantillon. Le sondage à deux degrés met en œuvre un double échantillonnage : sur les unités primaires et secondaires.

La probabilité d'inclusion d'ordre un d'un individu pour l'ensemble du plan de sondage est la probabilité de sélectionner l'unité primaire contenant cet individu multipliée par la probabilité de sélectionner l'individu dans cette unité primaire.

Les sondages à deux degrés possèdent les propriétés d'invariance et d'indépendances:

- L'invariance signifie que les sondages du deuxième degré ne dépendent pas de ce qui s'est passé au premier degré ;
- L'indépendance signifie que les tirages du deuxième degré sont indépendants les uns des autres (comme en stratification).

6.1.5. Échantillonnage en grappes

L'échantillonnage en grappes consiste à :

- subdiviser la population sous groupes ou grappes ;
- choisir au hasard un certain nombre de grappes ;
- prélever les observations chez tous les événements de chaque grappe.

6.2. Techniques d'échantillonnage non probabilistes

Les méthodes de sondages aléatoires qui ont été décrites précédemment supposent le tirage aléatoire de l'échantillon à partir d'une base de sondage, c'est-à-dire d'une liste exhaustive des individus composant la population étudiée. Lorsque de telles bases sont inexistantes ou indisponibles, lorsqu'il est trop coûteux de réaliser un sondage aléatoire, les méthodes dites non aléatoires, ou encore méthodes empiriques ou à choix raisonné peuvent être utilisées.

Il s'agit de :

- l'échantillonnage de commodité (de convenance, à l'aveuglette ou accidentel)
- l'échantillonnage volontaire
- l'échantillonnage au jugé
- l'échantillonnage par quotas
- l'échantillonnage Boule de neige

6.2.1. Échantillonnage de commodité (de convenance)

Parfois appelé échantillonnage à l'aveuglette ou accidentel, cet échantillonnage n'est pas normalement représentatif de la population cible, parce que ne sélectionnons des unités d'échantillonnage dans son cas que si nous pouvons y avoir facilement et commodément accès. Celui qui fait l'échantillonnage à l'aveuglette présume que la population est homogène : si les unités de la population sont toutes semblables, n'importe quelle unité peut être choisie pour l'échantillon. Il s'agit d'un échantillon d'individus qui se trouvaient accidentellement à l'endroit et au moment où l'information a été collectée. Les échantillons accidentels ne peuvent être considérés représentatifs d'aucune population. L'avantage évident de la méthode est qu'elle est facile à utiliser, mais la présence de biais annule énormément ce dernier. Même si ses applications utiles sont limitées, la technique peut donner des résultats exacts lorsque la population est homogène.

6.2.2. Échantillonnage volontaire

Cette méthode fait appel à des répondants volontaires. Elle sert souvent à sélectionner des particuliers pour des groupes de discussions approfondis, c'est-à-dire une mise à l'essai qualitative qui exclut la généralisation appliquée à la population complète. Le fait d'échantillonner des participants volontaires plutôt que la population en général peut introduire des biais. Souvent, à l'occasion des sondages d'opinion, seuls les gens qui se soucient assez fortement d'une façon ou d'une autre de la question étudiée ont tendance à y répondre. La majorité silencieuse n'y répond généralement pas, ce qui entraine un important biais sur le plan de la sélection.

6.2.3. Échantillonnage au jugé

La méthode d'échantillonnage au jugé est utilisée lorsque nous prélevons un échantillon en se fondant sur certains jugements au sujet de l'ensemble de la population. L'hypothèse qui sous-tend son utilisation est que l'enquêteur sélectionnera des unités qui seront les caractéristiques de la population. La méthode consiste à sélectionner des individus dont nous pensons avant de les interroger, qu'ils peuvent détenir l'information.

Les statisticiens utilisent souvent cette méthode dans le cadre d'études préparatoires comme des tests préalables de questionnaires et des discussions en groupe. La réduction du coût et du temps qu'exige l'acquisition de l'échantillon est l'un des avantages de l'échantillonnage au jugé. Le risque de ce type d'échantillonnage est de considérer des individus, apparemment représentatifs de la population étudiée.

6.2.4.Échantillonnage par quotas

L'échantillonnage par Quotas est l'une des formes les plus courantes d'échantillonnage non probabiliste. Il s'effectue jusqu'à ce qu'un nombre précis d'unités (quotas) pour diverses sous-populations ait été sélectionné. Puisqu'il n'existe aucune règle qui régirait la façon dont il faudrait s'y prendre pour remplir ces quotas, l'échantillonnage par quotas est réellement un moyen de satisfaire aux objectifs en matière de taille d'échantillon pour certaines sous-populations.

L'échantillonnage par quotas ressemble à l'échantillonnage stratifié parce que les unités semblables sont regroupées; la méthode de sélection des unités est cependant différente. Les unités sont sélectionnées aléatoirement dans l'échantillonnage probabiliste mais dans l'échantillonnage par quotas, une méthode non aléatoire est appliquée, c'est-à-dire les unités sollicitées qui ne sont pas disposées à participer sont remplacées par d'autres qui le sont, et nous ignorons en fait le biais de non-réponse.

La précision des estimateurs par quotas n'est pas calculable, puisqu'aucune probabilité n'est connue. En tenant compte du résultat numérique de précision, nous pouvons utiliser la formule de la variance d'un sondage stratifié, assimilant à une strate chaque sous-population sur laquelle nous devons respecter un quota.

6.2.5. Échantillonnage « Boule de neige »

Cette méthode est réservée aux populations composées d'individus dont l'identification est difficile ou qui possèdent des caractéristiques rares. La méthode consiste à faire construire l'échantillon par les individus eux-mêmes. Il suffit d'en identifier un petit nombre initial et de leur demander de faire appel à d'autres individus possédant les mêmes caractéristiques.

6.3. Estimation par intervalle de confiance

Il existe deux types d'estimation :

- l'estimation ponctuelle : dans le cas d'un échantillon aléatoire simple (EAS), le paramètre de population est estimé par la statistique calculée dans l'échantillon ;
- l'estimation par intervalle de confiance : l'analyste n'a pas besoin de connaître la vraie valeur du paramètre de population mais la fourchette dans laquelle se paramètre se trouve.

Si l'échantillon est aléatoire simple (EAS), alors le paramètre de population est estimé par la statistique calculée dans l'échantillon. Par exemple, dans une population dont la taille de N=1000, un échantillon aléatoire simple de taille n=10 est tiré et pour chaque individu de l'échantillon, la taille est mesurée. Dans ce cas, si la taille moyenne dans l'échantillon est de $\overline{x} = 1,50$ m, alors la taille moyenne non observée dans la population est estimée à $\hat{\mu} = 1,50$m.

La moyenne de la distribution échantillonnée est :

$$E\left(\hat{\theta}\right) = \frac{1}{T}\sum_{i=1}^{T}\hat{\theta}_i \qquad (6.6)$$

La variance de la distribution échantillonnée est :

$$Var(\hat{\theta})=\frac{1}{T}\sum_{i=1}^{T}\left[\hat{\theta}_i-E(\hat{\theta})\right]^2 \tag{6.7}$$

Soient :

θ un paramètre de population

$\hat{\theta}$ un estimateur de θ

$\hat{\theta}_i$ les estimations obtenues avec l'estimateur $\hat{\theta}$

T la taille de l'échantillon

Alors :

$$E(\hat{\theta})=\frac{1}{T}\sum_{i=1}^{T}\hat{\theta}_i \tag{6.8}$$

L'intervalle de confiance est construit sur base de l'estimation ponctuelle obtenue pour le paramètre étudié, la déviation standard de la distribution de l'échantillonnage (erreur standard) et le degré de confiance [20] :

$$IC_{95\%}(\theta)=\hat{\theta}_0 \pm t_{n-1;1-\frac{\alpha}{2}}\sqrt{Var(\hat{\theta})} \tag{6.9}$$

où t_{n-1} est quantile de Student à $(n-1)$ ddl, α le seuil de décision et $1-\alpha$ le niveau de confiance.

$\sqrt{Var(\hat{\theta})}$ est la racine carrée de la variance de la distribution échantillonnée.

Lorsque la taille de l'échantillon est grand (n>30) avec un coefficient de risque de 5 %, l'intervalle de confiance s'écrit :

$$IC(\theta)=\hat{\theta}_0 \pm 1{,}96\times ES(\hat{\theta}) \tag{6.10}$$

où ES est l'erreur standard, θ un paramètre de population, $\hat{\theta}$ un estimateur de θ et $\hat{\theta}_0$ une estimation ponctuelle de θ et $Var(\hat{\theta})$ La variance de la distribution échantillonnée.

6.3.1. Intervalle de confiance pour une moyenne

La moyenne $\bar{x}$ est calculée à partir d'un échantillon de taille *n* tiré dans une population de taille *N*. La distribution d'échantillonnage des moyennes a montré que la variable aléatoire $\overline{X}_n$ est une nouvelle variable aléatoire qui suit une loi normale de moyenne μ et d'écart-type $\frac{\sigma^2}{n}$ si l'échantillon est tiré avec remise. Dans le cas où il n'y a pas de remise, la variance de la distribution d'échantillonnage des moyennes est multipliée par le coefficient d'exhaustivité et vaut $\frac{\sigma^2}{n}\frac{N-n}{N-1}$. Bornons-nous au cas de tirage avec remise.

1) Si σ est connu, alors $\bar{X}_n \sim N\left(\mu,\frac{\sigma}{\sqrt{n}}\right)$. La standardisation montre que

$$U=\frac{\bar{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1) \tag{6.11}$$

Ainsi,

$$p\left(-z_{1-\frac{\alpha}{2}}\le U\le z_{1-\frac{\alpha}{2}}\right)=1-\alpha$$

$$\Leftrightarrow p\left(-z_{1-\frac{\alpha}{2}}\le \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\le z_{1-\frac{\alpha}{2}}\right)=1-\alpha$$

$$\Leftrightarrow p\left(-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\le \bar{x}-\mu\le z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

$$\Leftrightarrow p\left(-\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\le -\mu\le -\bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

$$\Leftrightarrow p\left(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\le \mu\le \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

L'intervalle de confiance de niveau $1-\alpha$ pour la moyenne μ lorsque σ est connu est :

$$IC_{(1-\alpha)100\,\%}(\mu)=\left[\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}};\bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right] \tag{6.12}$$

2) Si σ est inconnu, alors $\bar{X}_n \sim N\left(\mu,\frac{S_x}{\sqrt{n}}\right)$.

3) La standardisation donne :

$$T=\frac{\bar{X}_n-\mu}{\frac{S}{\sqrt{n}}}=\frac{\bar{X}_n-\mu}{\sqrt{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}}=\frac{\frac{\bar{X}_n-\mu}{\sigma}}{\sqrt{\frac{\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2}{n}}}\sim t(n-1) \tag{6.13}$$

T est de la forme $\dfrac{U}{\sqrt{\dfrac{V}{n}}}$ avec $U \sim N(0,1)$ et $V \sim \chi^2_{n-1}$.

Par le même truchement, calculons :

$$\begin{aligned}
& p\left(-t_{n-1;1-\frac{\alpha}{2}} \leq U \leq t_{n-11-\frac{\alpha}{2}}\right) = 1-\alpha \\
\Leftrightarrow\ & p\left(-t_{n-1;1-\frac{\alpha}{2}} \leq \frac{\overline{x}-\mu}{\frac{s}{\sqrt{n}}} \leq t_{n-11-\frac{\alpha}{2}}\right) = 1-\alpha \\
\Leftrightarrow\ & p\left(-t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \overline{x}-\mu \leq t_{n-11-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) = 1-\alpha \\
\Leftrightarrow\ & p\left(-\overline{x}-t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq -\mu \leq -\overline{x}+t_{n-11-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) = 1-\alpha \\
\Leftrightarrow\ & p\left(\overline{x}-t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \overline{x}+t_{n-11-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) = 1-\alpha
\end{aligned} \tag{6.14}$$

où $t_{n-1;1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi de Student à (n – 1) degrés de liberté.

Il vient que l'intervalle de confiance de la moyenne avec un niveau de confiance de $1-\alpha$ lorsque σ n'est pas connu est :

$$IC_{(1-\alpha)100\,\%}(\mu) = \left[\overline{x}-t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} ; \overline{x}+t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right] \tag{6.15}$$

Remarque : Si $n > 30$ alors $t_{n-1;1-\frac{\alpha}{2}} = z_{1-\frac{\alpha}{2}}$.

Pour une moyenne, l'intervalle de confiance à 95 % s'écrit :

$$IC_{95\%}(\mu) = \overline{x} \pm t_{n-1;1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \tag{6.16}$$

ce qui veut dire que $\mu \in \left[\overline{x}-1{,}96\frac{\sigma}{\sqrt{n}} ; \overline{x}+1{,}96\frac{\sigma}{\sqrt{n}}\right]$.

6.3.2. Intervalle de confiance pour une proportion

L'estimateur non biaisé de la proportion p est :

$$F_n = \frac{X}{n} \text{ avec } X \sim Bin(n, p) \tag{6.17}$$

Déterminons les proportions $f_1 = \frac{x_1}{n}$ et $f_2 = \frac{x_2}{n}$ avec $p(X \leq x_1) = \frac{\alpha}{2}$ et $p(X \geq x_2) = \frac{\alpha}{2}$.

L'intervalle de confiance de niveau de confiance $1-\alpha$ autour de la vraie proportion est $[f_1, f_2]$.

Remarque : Lorsque la taille de l'échantillon n dépasse 30, la loi binomiale est approximée par la loi normale :

$$F_n \sim N\left(\text{p}, \sqrt{\frac{p(1-p)}{n}}\right) \tag{6.18}$$

La standardisation donne :

$$T = \frac{F_n - \text{p}}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1) \tag{6.19}$$

Ainsi,

$$p\left(-z_{1-\frac{\alpha}{2}} \leq T \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Leftrightarrow p\left(-z_{1-\frac{\alpha}{2}} \leq \frac{F_n - \text{p}}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Leftrightarrow p\left(-z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \leq F_n - \text{p} \leq z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right) = 1-\alpha$$

$$\Leftrightarrow p\left(-F_n - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \leq -p \leq -F_n + z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right) = 1-\alpha$$

$$\Leftrightarrow p\left(F_n - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \leq p \leq F_n + z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right) = 1-\alpha$$

Il vient que l'intervalle de confiance de la proportion avec un niveau de confiance de $1-\alpha$ est :

$$IC_{(1-\alpha)100\%}(\pi)=\left[p-z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}};\mathrm{p}+z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right] \tag{6.20}$$

Pour une proportion, cet intervalle devient [20] :

$$IC_{95\%}(\pi)=p\pm t_{n-1;1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \tag{6.21}$$

6.3.3. Intervalle de confiance pour une variance pour une population normale

1) Si la moyenne μ de la population est connue, l'estimateur de σ^2 est :

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \tag{6.22}$$

Sachant que

$$\frac{\hat{\sigma}^2}{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2=\frac{1}{n}\sum_{i=1}^{n}z_i^2$$

il vient que $\frac{n\hat{\sigma}^2}{\sigma^2}\sim\chi^2(n)$.

Ainsi,

$$p\left(\chi^2_{\frac{\alpha}{2},n}\le\frac{n\hat{\sigma}^2}{\sigma^2}\le\chi^2_{1-\frac{\alpha}{2},n}\right)=1-\alpha \tag{6.23}$$

Il vient

$$p\left(\frac{n\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2},n}}\le\sigma^2\le\frac{n\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2},n}}\right)=1-\alpha \tag{6.24}$$

L'intervalle de confiance de σ^2 lorsque μ est connu est :

$$IC(\sigma^2)=\left[\frac{n\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2},n}},\frac{n\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2},n}}\right] \tag{6.25}$$

où $\chi^2_{1-\frac{\alpha}{2},n}$ et $\chi^2_{\frac{\alpha}{2},n}$ sont les quantiles d'ordre $1-\frac{\alpha}{2}$ et $\frac{\alpha}{2}$ de la loi du chi-deux à n degrés de liberté.

2) Si la moyenne μ de la population est inconnue, l'estimateur non biaisé de variance est

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \tag{6.26}$$

La division membre à membre par σ^2 donne :

$$\frac{S^2}{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{\sigma}\right)^2 \tag{6.27}$$

Il vient

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{\sigma}\right)^2 \sim \chi^2(n-1) \tag{6.28}$$

L'intervalle de confiance de σ^2 lorsque μ est non connu est :

$$IC\left(\sigma^2\right) = \left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}\right] \tag{6.29}$$

ou encore

$$IC\left(\sigma^2\right) = \left[\frac{nS^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}, \frac{nS^2}{\chi^2_{\frac{\alpha}{2},n-1}}\right] \tag{6.30}$$

où $\chi^2_{1-\frac{\alpha}{2},n-1}$ et $\chi^2_{\frac{\alpha}{2},n-1}$ sont les quantiles d'ordre $1-\frac{\alpha}{2}$ et $\frac{\alpha}{2}$ de la loi de chi-deux à $n-1$ degrés de liberté.

L'intervalle de confiance de variance pour la variance est donné par :

$$p\left(\frac{(n-1)\mathrm{S}_x^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)\mathrm{S}_x^2}{\chi^2_{\frac{\alpha}{2}}}\right) = 1-\alpha \tag{6.31}$$

Remarque :

Si n > 30, alors $\chi^2_{\alpha,n-1} \approx \frac{1}{2}\left(t_\alpha + \sqrt{2n-3}\right)^2$ si bien que :

$$IC\left(\sigma^2\right) = \left[\frac{2(n-1)S^2}{\left(t_{1-\frac{\alpha}{2}} + \sqrt{2n-3}\right)^2}, \frac{2(n-1)S^2}{\left(t_{\frac{\alpha}{2}} + \sqrt{2n-3}\right)^2}\right] \tag{6.32}$$

La symétrie de la loi normale centrée-réduite assure que $t_{\frac{\alpha}{2}} = -t_{1-\frac{\alpha}{2}}$.

L'intervalle de confiance de l'écart-type est donné par [21,22] :

$$p\left(\sqrt{\frac{(n-1)S_x^2}{\chi^2_{1-\frac{\alpha}{2}}}} < \sigma < \sqrt{\frac{(n-1)S_x^2}{\chi^2_{\frac{\alpha}{2}}}}\right) = 1-\alpha \tag{6.33}$$

6.3.4. Intervalle de confiance pour une médiane

L'erreur standard de la médiane est :

$$S_{\mu_e} = S_x\sqrt{\frac{\pi}{2n}} \tag{6.34}$$

L'intervalle de confiance de la médiane est donné par :

$$p\left(\mu_e - z_{\frac{\alpha}{2}} S_x\sqrt{\frac{\pi}{2n}} < mediane < \mu_e + z_{\frac{\alpha}{2}} S_x\sqrt{\frac{\pi}{2n}}\right) = 1-\alpha \tag{6.35}$$

Si un échantillon est extrait d'une population approximativement normale et si son effectif est relativement grand $(n > 60)$, la distribution d'échantillonnage de la médiane s'approche de la loi normale.

Exemple 1 :

La moyenne du taux de cholestérol d'une certaine population d'hommes en bonne santé vaut 240mg/dl avec un écart-type de 40mg/dl. Un chercheur envisage de comparer les taux de cholestérol d'un échantillon aléatoire de 100 hommes ayant subi un pontage coronarien l'année précédente. La moyenne de taux de cholestérol de cet échantillon vaut 260mg/dl.

a) Est-il raisonnable de dire que le taux moyen de cholestérol des hommes ayant subi un pontage coronarien est plus élevé que le taux moyen de cholestérol de la population des hommes en bonne santé ? Vérifiez-le à l'aide d'intervalle de 95 %.

b) S'il avait obtenu les mêmes résultats mais en étudiant un échantillon de plus petite taille $(n=9)$, aurait-il pu conclure ?

Solution

Soit x la variable qui désigne le taux de cholestérol. Les données sont :

	Population en bonne santé	Population ayant fait le pontage coronarien
Moyenne	$\mu_0 = 240$	$\bar{x} = 260$
Écart-type	$\sigma_0 = 40$	-
Taille	-	$n = 100$

L'intervalle de confiance à 95 % du taux moyen de cholestérol dans la population vaut :

$$\begin{aligned} IC_{95\%}(\mu) &= \bar{x} \pm t_{n-1;1-\frac{\alpha}{2}} \frac{\sigma_0}{\sqrt{n}} \\ &= 260 \pm 1,96 \times \frac{40}{10} \\ &= 260 \pm 7,84 \\ &= [252,16; 267,84] \end{aligned}$$

Avec un niveau de confiance de 95 %, le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien est compris entre 252,16mg/dl et 267,84mg/dl. Cet intervalle de confiance ne contient pas $\mu_0 = 240$

Conclusion :

Il existe une différence significative entre le taux moyen de cholestérol dans la population d'hommes en bonne santé et le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien.

S'il avait obtenu les mêmes résultats mais en étudiant un échantillon de plus petite taille, aurait-il pu conclure ?

Solution

$n = 9$

L'intervalle de confiance à 95 % du taux moyen de cholestérol dans la population est :

$$\begin{aligned} IC_{95\%}(\mu) &= \bar{x} \pm t_{n-1;1-\frac{\alpha}{2}} \frac{\sigma_0}{\sqrt{n}} \\ &= 260 \pm 2,306 \times \frac{40}{\sqrt{9}} \\ &= 260 \pm 30,75 \\ &= [229,25; 290,75] \end{aligned}$$

Avec un niveau de confiance de 95 %, le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien est compris entre 229,25mg/dl et 290,75 mg/dl. Cet intervalle de confiance contient la valeur 240.

Conclusion :

Il n'existe pas de différence significative entre le taux moyen de cholestérol dans la population d'hommes en bonne santé et le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien.

Exemple 2

Lors d'enquête sur la durée du sommeil des enfants de 2 à 3 ans effectuée sur un échantillon de 540 enfants d'un département français, on a trouvé 86 enfants présentant des troubles de sommeil. Cette proportion est-elle significativement différente de la proportion de 15 % théoriquement attendue ? Construire un intervalle de confiance à 99 % de la proportion d'enfants présentant les troubles de sommeil dans ce département.

Solution

$n = 540$

La proportion d'enfants présentant des troubles de sommeil vaut :

$$p = \frac{o_1}{n} = \frac{86}{540}$$

La proportion d'enfants ne présentant pas les troubles de sommeil vaut :

$$q = \frac{o_2}{n} = \frac{n - o_1}{n} = \frac{454}{540}$$

L'intervalle de confiance à 99 % de la proportion d'enfants présentant des troubles de sommeil dans le département est :

$$
\begin{aligned}
IC_{99\%}(\pi) &= p \pm t_{n-1;1-\frac{\alpha}{2}} \sqrt{\frac{pq}{n}} \\
&= \frac{86}{540} \pm 2,576 \sqrt{\frac{\frac{86}{540} \times \frac{454}{540}}{540}} \\
&= 0,159 \pm 0,041 \\
&= [0,118; 0,200]
\end{aligned}
$$

Vérifions les conditions :

$$
\underbrace{\begin{aligned}
&nBI = 540(0,118) \simeq 64 > 5 \\
&nBS = 540(0,200) = 108 > 5 \\
&n(1-B) = 540(1-0,118) \simeq 476 > 5 \\
&n(1-BS) = 540(1-0,200) = 432 > 5
\end{aligned}}_{\substack{4\,\text{conditons} \\ IC\ \text{valide}}}
$$

Avec un niveau de confiance de 99 %, la proportion d'enfants présentant du trouble de sommeil dans le département est comprise entre 11,8 % et 20 %. Cet intervalle de confiance contient 15 %. Donc, il n'y a pas de raison de dire que la proportion d'enfants présentant des troubles de sommeil dans le département est significativement différente de 15 %.

Soit une population obéissant à une loi normale de moyenne μ (inconnue) et d'écart-type σ (inconnu).Pour $v = n-1$ degrés de liberté.

6.4. Tests d'hypothèses

L'objectif des tests statistiques ou tests d'hypothèses est de répondre à des questions suivantes tout en utilisant des affirmations (non gratuites, non hâtives) basées sur une probabilité de se tromper :

- L'entreprise A est-elle plus performante que l'entreprise B ?
- Les étudiants de l'École Normale Supérieure se trouvent plus un emploi que ceux de l'Université du Burundi ?
- Les femmes urbaines sont-elles plus obèses que les femmes rurales ?
- Les étudiants de la Section Mathématique réussissent-ils mieux en Statistique et probabilités que les étudiants de la Section Physique-Technologie ?

Les éléments de réponses à ces questions proviennent des statistiques calculées dans l'échantillon entre autres la taille de l'échantillon, la moyenne, la déviation standard et la proportion.

6.4.1. Principe

Le cadre théorique de l'échantillonnage sur lequel se basent les théories d'inférence statistique classique est celui d'un échantillonnage aléatoire simple (EAS) effectué à partir de la population infinie.

Un test d'hypothèses comporte 3 étapes [21,22] :

a) poser une hypothèse au niveau des paramètres de la population

b) sous l'hypothèse nulle, calculer un test à partir des observations faites dans le/les échantillon(s).

c) à partir du résultat du test, tirer une conclusion par rapport à l'hypothèse nulle posée.

L'hypothèse testée s'appelle **hypothèse nulle** et classiquement notée H_0.

Le nom donné à cette hypothèse vient du fait que, même si ce souvent des effets qui doivent être mis en évidence, l'analyste teste l'absence d'effet : l'égalité de deux moyens, de deux proportions.

Par exemple, H_0 se définit souvent de façon complémentaire et est définie en considérant qu'il n'y a pas de changements :

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_0 &: \pi_1 = \pi_2 \end{aligned} \tag{6.36.a}$$

ce qui est équivalent à :

$$\begin{aligned} H_0 &: \mu_1 - \mu_2 = 0 \\ H_0 &: \pi_1 - \pi_2 = 0 \end{aligned} \tag{6.36.b}$$

L'hypothèse nulle peut être :

- L'entreprise A est aussi performante que l'entreprise B
- Les étudiants de l'École Normale Supérieure se trouvent de la même manière un emploi que ceux de l'Université du Burundi
- Les femmes urbaines sont aussi obèses que les femmes rurales
- Les étudiants de la Section Mathématique réussissent au même titre en Statistique et probabilités que les étudiants de la Section Physique-Technologie

Réaliser un test statistique consiste à décider si une hypothèse H_0 concernant une ou plusieurs populations est confirmée ou non par des résultats d'échantillonnage. Les hypothèses statistiques sont d'un type particulier : elles ne peuvent être démontrées puisque la/les populations concernées ne sont pas enquêtées entièrement. Tout ce qui est à la disposition de l'analyste concerne les résultats d'expérience, d'observations réalisées dans des échantillons. Par contre, H_0 falsifiable : une expérience peut donner un résultat peu plausible avec une hypothèse nulle H_0. Les conclusions qui peuvent être tirées d'un test d'hypothèse ne sont pas définitives. Elles amèneront simplement à dire que les résultats observés sur les échantillons sont tels qu'ils n'apportent aucune preuve évidente contre H_0 ou, au contraire, que juste à un certain degré ils permettent de mettre sérieusement l'hypothèse nulle H_0 en doute.

Le test se calcule sur l'hypothèse nulle H_0 et diffère selon l'hypothèse nulle posée.

La décision du test est de forme dichotomique : rejet ou non rejet de l'hypothèse nulle.

L'hypothèse contraire à l'hypothèse nulle H_0 s'appelle **hypothèse alternative ou contraire** et se note H_1.

Par exemple, l'hypothèse alternative qui est l'hypothèse qui se veut être démontrée est :

$$\begin{aligned} &H_1 : \mu_1 \neq \mu_2 \\ &H_1 : \pi_1 \neq \pi_2 \end{aligned} \tag{6.37.a}$$

ce qui est équivalent à :

$$\begin{aligned} &H_1 : \mu_1 - \mu_2 \neq 0 \\ &H_1 : \pi_1 - \pi_2 \neq 0 \end{aligned} \tag{6.37.b}$$

L'hypothèse alternative peut être :

- L'entreprise A est plus performante que l'entreprise B
- Les étudiants de l'École Normale Supérieure se trouvent plus facilement un emploi que ceux de l'Université du Burundi
- Les femmes urbaines sont plus obèses que les femmes rurales
- Les étudiants de la Section Mathématique réussissent mieux en Statistique et probabilités que les étudiants de la Section Physique-Technologie

Le non rejet de H_0 ne signifie pas qu'elle est forcément vraie mais que les données obtenues ne sont pas suffisantes pour être en contradiction avec H_0 pour que l'hypothèse contraire ou alternative H_1 soit préférée. Les deux hypothèses H_0 et H_1 ne jouent pas le même rôle. Si l'une est vraie, alors l'autre est fausse et vice-versa. La construction d'un test nécessite la définition de deux hypothèses et d'un mécanisme de décision.

Le choix de l'une ou de l'autre de ces 2 décisions possibles (rejet ou non rejet de l'hypothèse nulle) est basé sur la probabilité d'observer, si l'hypothèse nulle H_0 est vraie, un résultat aussi éloigné ou plus éloigné de la valeur prévue sans l'hypothèse nulle H_0 que celui qui a été observé dans le/les échantillons. Si cette probabilité est suffisamment petite, alors les estimations sont peu plausibles avec l'hypothèse nulle H_0 et la décision sera de rejeter celle-ci. Au contraire, si cette probabilité est grande, alors les observations sont plausibles avec l'hypothèse nulle H_0 et la décision sera de ne pas rejeter celle-ci.

Le seuil définissant la probabilité suffisamment petite est fixé arbitrairement. Il est choisi à 5 % le plus souvent et dans certains cas, à 1 % ou à 1 %o. Il est compris entre 1 et 10 %. Pratiquement, la probabilité définie ci-dessus est équivalente à la probabilité d'observer, si l'hypothèse nulle est vraie, une valeur du test aussi grande ou plus grande que celle calculée à partir d'observations à étape précédente de la réalisation du test. Cette dernière est notée p et est appelée $p-$valeur ou $p-$value en anglais. Si la $p-$ value est inférieure au seuil de signification choisi, alors l'hypothèse nulle H_0 sera rejetée et le test sera dit **significatif**. Au contraire, si la $p-$value est supérieure ou égale au seuil de signification choisi, alors l'hypothèse nulle H_0 ne sera pas rejetée et le test sera dit **non significatif**.

6.4.2. Types d'erreurs

La réalisation d'un test implique une décision : rejet ou non rejet de l'hypothèse nulle. Cette décision base sur les observations faites dans le/les échantillons concernent des populations où la situation réelle est inconnue. C'est donc une décision qui risque d'être accompagnée d'erreurs. Deux types d'erreurs ou risques sont possibles en tirant la conclusion lors d'un test d'hypothèses :

- α appelé risque de 1ère espèce est la probabilité de se tromper en rejetant H_0 alors qu'elle est vraie.
- β appelé risque de 2ème espèce est la probabilité de se tromper en ne rejetant pas H_0 alors qu'elle est fausse.
- n est la taille de l'échantillon. Il faut choisir la plus grande taille possible mais elle est limitée par les conditions de l'expérience et le coût.

Le tableau 13 montre la décision, les hypothèses et les erreurs qui accompagnent chaque décision.

Tableau 13 : Types d'erreur

Hypothèse *Décision*	H_0 vraie	H_1 vraie
Rejet	Mauvaise de décision Erreur de 1ère espèce α	Bonne décision Puissance d'un test $1-\beta$
Non rejet	Bonne décision Niveau de confiance $1-\alpha$	Mauvaise décision Erreur de 2ème espèce β

6.5. Étapes de réalisation d'un test

1° Choix de H_0

Le choix de l'hypothèse nulle dépend du type de test à faire et donc de la question posée. Par exemple, pour les tests paramétriques (comparaison des moyennes, des variances, des modes, des proportions) pour lesquels la distribution du paramètre ou des observations doit être connue, l'hypothèse nulle porte sur un ou plusieurs paramètres θ de population. Pour les tests non paramétriques, l'hypothèse nulle est très générale. Pour les tests d'adéquation, l'hypothèse nulle porte sur la distribution de la variable. Par exemple, le taux de mauvais contrôle glycémique suit-il une loi normale ? La distribution empirique s'accorde-t-elle avec la distribution observée ? Pour le test d'indépendance ou d'homogénéité de deux variables aléatoires, l'hypothèse nulle dit qu'il y a indépendance ou homogénéité entre deux variables (indépendance entre le tabagisme et l'alcoolisme par exemple).

2° **Choix de H_1**

L'hypothèse alternative est une hypothèse qui sera favorisée si H_0 est rejetée. H_1 peut être hypothèse du genre $\theta_1 \neq \theta_2$ (test bilatéral) ou $\theta_1 > \theta_2$ (test unilatéral).

3° Choix du seuil de decision

Ce seuil est souvent choisi à 5 %.

4° Choix du test

Le test et la statistique de test sont choisis en fonction de l'objectif de l'étude. À cette étape de réalisation d'un test, la région critique appelée « région de rejet » est déterminée. La région critique R est une partie de l'ensemble des valeurs possibles de la variable de décision T (une variable aléatoire qui dépend de l'échantillon) telle que $\alpha = p(T \in R \mid H_0\, vraie)$. La règle de décision est que, si la valeur t de la variable aléatoire T appartient à la région de rejet R, alors H_0 est rejetée et si t n'appartient pas à R, c'est-à-dire $t \in \overline{R}$, alors H_0 n'est pas rejetée.

- Ne pas rejeter H_0 signifie que les calculs faits sur l'échantillon ne sont pas en contradiction significative avec l'hypothèse nulle (H_0).
- Rejeter H_0 signifie que t est considérée comme trop improbable sous H_0 pour que H_0 ne soit rejetée.
- α est la probabilité de rejeter à tort l'hypothèse H_0.

5° Calcul de la p-value

La p-value est calculée sur base de la valeur de la statistique de test et de la statistique tabulée avec les degrés de liberté.

5° Décision

Si la p-value est inférieure au seuil de décision, l'hypothèse nulle sera rejetée. La puissance du test est aussi calculée.

6° Conclusion

La conclusion est tirée en fonction de la question posée et de l'hypothèse nulle formulée mathématiquement et en français.

6.6. Principaux tests statistiques

6.6.1. Test sur une moyenne

Le test sur une moyenne consiste à tester l'égalité entre une moyenne et une moyenne théorique donnée.

Soient :

n la taille de l'échantillon

$\bar{x}$ la moyenne observée dans l'échantillon

s_x la déviation standard calculée dans l'échantillon

μ la moyenne de la population dont l'échantillon a été extrait

μ_0 la moyenne théorique donnée

Les hypothèses des tests sont :

$H_0 : \mu = \mu_0$ (Hypothèse nulle)

$H_0 : \mu \neq \mu_0$ (Hypothèse alternative) (6.38)

Il s'agit d'un test bilatéral. En supposant que H_0 soit vraie, la statistique de test est [16] :

$$t^{obs} = \frac{|\bar{x} - \mu_0|}{\frac{s}{\sqrt{n}}} \text{ à (n-1) ddl} \quad (6.39)$$

La $p-$value sera déterminée en utilisant la distribution t (distribution de Student) à $(n-1)ddl$. La décision portera sur le rejet H_0 (test significatif) si la $p-$value est inférieure ou égale au seuil de signification choisie et au non rejet H_0 (test non significatif) si la $p-$value est supérieure au seuil de signification choisie.

Exemple 1 :

La moyenne du taux de cholestérol d'une certaine population d'hommes en bonne santé vaut 240 mg/dl avec un écart-type de 40 mg/dl. Un chercheur envisage de comparer les taux de cholestérol d'un échantillon aléatoire de 100 hommes ayant subi un pontage coronarien l'année précédente. La moyenne du taux de cholestérol dans cet échantillon vaut 260 mg/dl. Au seuil de 5 %, la moyenne du taux du taux de cholestérol chez les hommes ayant subi un pontage coronarien l'année précédente est-il significativement différent du taux moyen de cholestérol des hommes en bonne santé ? Vérifiez-le à l'aide d'un test statistique.

Solution

$\mu_0 = 240$

$\sigma_0 = 40$

$n = 100$

$\bar{x} = 260$

Il s'agit d'un test t de Student à un échantillon (comparaison d'une moyenne à une moyenne théorique)

Les hypothèses de test sont :

$H_0 : \mu = 240$

Le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien de 240 mg/dl

$H_1 : \mu \neq 240$

Le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien est différent de 240mg/dl.

Seuil la décision $\alpha = 5\% = 0{,}05$

La statistique de test vaut :

$$t^{obs} = \frac{|\bar{x} - \mu_0|}{\frac{s_x}{\sqrt{n}}} = \frac{|260 - 240|}{\frac{40}{\sqrt{100}}} = \frac{20}{4} = 5 \text{ à 99 ddl}$$

$t^{obs} > 3{,}291$

$p-$value<0,001<0, 05

$p-$value<0, 05

Décision : Rejet de H_0(test significatif)

Conclusion: Au seuil de 5 %, le taux moyen de cholestérol dans la population d'hommes ayant subi un pontage coronarien est significativement différent de 240 mg/dl.

Exemple 2 :

Dans un échantillon de 14 filles nées dans une maternité d'un pays en voie de développement, on observe un poids moyen de naissance égal à 2,81 kg avec une déviation standard de 0,44 kg. Le poids moyen de naissance dans une population de référence de filles vaut 3,2 kg. Le poids moyen observé dans un échantillon de 14 filles est-il compatible avec cette moyenne théorique au seuil de 5 % ? Vérifiez-le à l'aide d'un test statistique puis à l'aide d'un intervalle de confiance à 95 %.

Solution

$n = 14$

$\bar{x} = 2,81$

$s_x = 0,44$

$\mu_0 = 3,20$

Seuil : $\alpha = 5\% = 0,05$

Il s'agit d'un test *t* de Student à un échantillon (comparaison d'une moyenne à une autre moyenne théorique).

Les hypothèses de test sont :

$$H_0 : \mu = 3,20$$

Cette hypothèse veut dire que le poids moyen des naissances des filles dans la population vaut 3,2 kg.

$$H_1 : \mu \neq 3,20$$

Cette hypothèse signifie que le poids moyen des naissances des filles dans la population est différent de 3,2 kg.

La statistique de test vaut :

$$t^{obs} = \frac{|\bar{x} - \mu_0|}{\frac{s}{\sqrt{n}}} = \frac{|2,81 - 3,20|}{\frac{0,44}{\sqrt{14}}} = 3,32 \text{ à } 13 \, ddl$$

0,001<p-value<0,01<0,05

Décision : Rejet de H_0 (test significatif)

Conclusion : Au seuil de 5 %, le poids moyen des naissances des filles dans la population d'où les 14 filles ont été tirées est significativement différent de 3,2 kg.

L'intervalle de confiance à 95 % du poids moyen des naissances des filles dans la population d'où les 14 filles ont été tirées est :

$$\begin{aligned} IC_{95\%}(\mu) &= \bar{x} \pm t_{n-1;0,975}\frac{s}{\sqrt{n}} \\ &= 2,81 \pm 2,160 \times \frac{0,44}{\sqrt{14}} \\ &= 2,81 \pm 0,254 \\ &= [2,56;3,06] \end{aligned}$$

Avec un niveau de confiance de 95 %, le poids moyen des naissances des filles dans la population d'où les 14 filles ont été tirées est compris entre 2,56 kg et 3,06 kg. Cet intervalle ne contient pas la valeur 3,2 (il est significatif).

Décision : Rejet de H_0 (IC significatif)

Conclusion : Au seuil de 5 %, le poids moyen des naissances des filles dans la population d'où les 14 filles ont été tirées est significativement différent de 3,2 kg. Le poids moyen observé dans l'échantillon de 14 filles est donc incompatible avec la moyenne théorique de 3,2 kg au seuil de 5 %.

6.6.2. Test sur une proportion

Soient :

n la taille de l'échantillon

o_1 le nombre de sujets ayant la caractéristique considérée

$o_2 = n - o_1$ le nombre de sujets sans caractéristique considérée.

p la proportion observée dans l'échantillon, soit $\frac{o_1}{n}$

π la proportion dans la population d'où l'échantillon a été tiré

π_0 La proportion théoriquement donnée.

Les hypothèses de test sont :

$H_0 : \pi = \pi_0$ (hypothèse nulle)

$H_1 : \pi \neq \pi_0$ (hypothèse alternative) (6.40)

En supposant que H_0 soit vraie, la statistique du test de Student est [16] :

$$t^{obs} = \frac{|p-\pi_0|}{\sqrt{\frac{p(1-p)}{n}}} \text{ à (n-1) ddl} \tag{6.41}$$

La statistique du chi-carré est :

$$\chi^2_{obs} = \frac{(o_1-e_1)^2}{e_1} + \frac{(o_2-e_2)^2}{e_2} \text{ à 1 ddl} \tag{6.42}$$

avec e_1 et e_2 des nombres (effectifs) attendus de sujets sous H_0 :

$$e_1 = n\pi_0 \tag{6.43}$$

et

$$e_2 = n(1-\pi_0) \tag{6.44}$$

La $p-$value sera déterminée en utilisant la distribution t à *(n-1)* ddl ou la distribution χ^2 à 1 ddl. La décision portera au rejet de H_0 (test significatif) si la $p-$value est inférieure ou égale à α (seul de décision) et au non rejet de H_0 (test non significatif) si $p-$valu est supérieur à α .

Exemple

Lors d'une enquête sur la durée du sommeil des enfants de 2à3ans effectué sur un échantillon de 540 enfants d'un département Français, on a trouvé 86 enfants présentant des troubles de sommeil. Est-il raisonnable de dire que la proportion vraie d'enfants présentant des troubles de sommeil dans ce département est différente de la proportion de 15 % théoriquement attendue au seuil de $\alpha = 1\%$? Vérifiez-le à l'aide de :

- Un test *t* de Student
- Un intervalle de confiance à 99 %
- Un chi-deux (chi-carré)

Solution

La proportion d'enfants présentant des troubles de sommeil dans l'échantillon est :

$$p = \frac{86}{540} = \frac{43}{270}$$

La proportion d'enfants ne présentant pas des troubles de sommeil dans l'échantillon est :

$$q = 1 - p = 1 - \frac{86}{540} = 1 - \frac{43}{270} = \frac{227}{270}$$

Les hypothèses de test sont :

$$H_0 : \pi = 0,15$$

La proportion d'enfants présentant des troubles de sommeil dans le département vaut 15 %.

$$H_1 : \pi \neq 0,15$$

La proportion d'enfants présentant des troubles de sommeil dans le département est différente de 15 %.

Il s'agit d'un test t de Student sur un échantillon (comparaison d'une proportion à une proportion théorique).

Seuil de décision : $\alpha = 1\% = 0,01$

La statistique du test de Student est [16] :

$$t^{obs} = \frac{|p - \pi_0|}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{\left|\frac{43}{270} - 0,15\right|}{\sqrt{\frac{0,15(0,85)}{540}}} \simeq 0,60 \text{ à } 339\, ddl$$

0 ,01<0,50<p-value<0,90 $\Rightarrow$ P-value >0,01

Décision : Non rejet de H_0 (test non significatif)

Conclusion : Au seuil de 1 %, la proportion d'enfants présentant des troubles de sommeil dans le Département n'est pas significativement différente de 15 %.

La statistique du chi-carré est :

$$\chi^2_{obs} = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} \text{ à } 1\, ddl$$

avec

$$e_1 = n\pi_0 = 540 \times 0,15 = 81$$
$$e_2 = n(1 - \pi_0) = 540 \times 0,85 = 459$$

Il vient :

$$\chi^2_{obs} = \frac{(86-81)^2}{81} + \frac{(454-459)^2}{459} = \frac{25}{81} + \frac{25}{459} = 0,36 \text{ à } 1\, ddl$$

0 ,01<0,50<p-value<0,90

$\Rightarrow$ P-value >0,01

Décision : Non rejet de H_0 (test non significatif)

Conclusion : Au seuil de 1 %, la proportion d'enfants présentant des troubles de sommeil dans le Département n'est pas significativement différente de 15 %.

L'intervalle de confiance à 99 % de la proportion vraie d'enfants présentant des troubles de sommeil dans le Département est :

$$
\begin{aligned}
IC_{95\%}(\pi) &= p \pm t_{n-1}\sqrt{\frac{p(1-p)}{n}} \\
&= \frac{43}{270} \pm 2,576\sqrt{\frac{\frac{43}{270}\left(1-\frac{43}{270}\right)}{540}} \\
&= 0,159 \pm 0,041 \\
&= [0,119;0,200]
\end{aligned}
$$

Comme

$$
\begin{cases}
N\times BI=540\times 0,119=64,26>5 \\
N\times(1\text{-}BI)=540\times 0,881=475,74>5 \\
N\times BS=540\times 0,200=108>5 \\
N\times(1\text{-}BS)=540\times 0,800=432>5
\end{cases}
\Rightarrow IC\ valide
$$

Décision : Non rejet de H_0 (IC non significatif)

Conclusion : Au seuil de 1 %, la proportion d'enfants présentant des troubles de sommeil dans le Département n'est pas significativement différente de 15 %.

6.6.3. Test sur deux moyennes

Soient :

n_1 la taille du premier échantillon

n_2 la taille du deuxième échantillon

$\overline{x}_1$ la moyenne de la variable dans le premier échantillon

$\overline{x}_2$ la moyenne de la variable dans le deuxième échantillon

s_1^2 la variance de la variable dans le premier échantillon

s_2^2 la variance de la variable dans le deuxième échantillon

s_p^2 la variance poolée (commune)

s_p l'écart-type poolé (commun)

μ_1 la moyenne de la variable dans la première population

μ_2 la moyenne de la variable dans la deuxième population

σ_1^2 la variance de la variable dans la première population

σ_2^2 la variance de la variable dans la deuxième population

Les hypothèses du test d'égalité des variances sont :

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ (Les variances sont égales)} \qquad (6.45.a)$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2 \text{ (Les variances sont inégales)} \qquad (6.45.b)$$

La comparaison des variances est effectuée à l'aide du test de BARTELET basé sur le test de Fisher.

La statistique de test est :

$$F^{obs} = \frac{S_{max}^2}{S_{min}^2} \qquad (6.46)$$

Si $F^{obs} \leq 3$, les variances peuvent être considérées comme homogènes.

Si $F^{obs} > 3$, les variances peuvent être considérées comme non homogènes.

La variance poolée (si les variances sont égales) vaut :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad (6.47)$$

Les hypothèses du test d'égalité des moyennes sont :

$$H_0 : \mu_1 = \mu_2 \text{ ou } H_0 : \mu_1 - \mu_2 = 0 \text{ (Les moyennes sont égales)} \qquad (6.48)$$

$$H_1 : \mu_1 \neq \mu_2 \text{ ou } H_1 : \mu_1 - \mu_2 \neq 0 \text{ (Les moyennes sont différentes)} \qquad (6.49)$$

En supposant que H_0 soit vraie et que les variances soient égales, la statistique de test est [16] :

$$t^{obs} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ à } (n_1 + n_2 - 2)\, ddl \qquad (6.50)$$

On conclura au rejet de H_0 (test significatif) si la $p-$value est inférieure ou égale au seuil de signification choisie et au non-rejet de H_0 (test non significatif) si la $p-$value est supérieure au seuil de signification choisie.

L'intervalle de confiance à 95 % de la différence entre 2 moyennes est donné par :

$$IC_{95\%}\left(\mu_1-\mu_2\right)=\left|\overline{x}_1-\overline{x}_2\right|\pm t_{n_1+n_2-2;0,975}\,s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \tag{6.51}$$

Si les variances ne sont pas supposées égales et en supposant que H_0 soit vraie, la statistique du test de Welch est :

$$t^{obs}=\frac{\left|\overline{x}_1-\overline{x}_2\right|}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}\text{ à }\upsilon\,ddl \tag{6.52}$$

avec le nombre de degrés de liberté donnés calculé par Welch et Satterwhite [29] :

$$\upsilon=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2\left(n_1-1\right)}+\frac{s_2^4}{n_2^2\left(n_2-1\right)}} \tag{6.53}$$

Une fois ces degrés de liberté calculés, la table de la loi de Student est utilisée pour lire la p-value.

La conclusion portera sur le rejet de H_0 (test significatif) si la p −value est inférieure ou égale au

seuil de signification choisie et au non rejet de H_0 (test non significatif) si la p −value est supérieure au seuil de signification choisie.

L'intervalle de confiance à 95 % de la différence entre 2 moyennes est donné par :

$$IC_{95\%}\left(\mu_1-\mu_2\right)=\left|\overline{x}_1-\overline{x}_2\right|\pm t_{n_1+n_2-2;0,975}\sqrt{\frac{s_1^2}{n_1}+\frac{s_1^2}{n_2}} \tag{6.54}$$

Si les échantillons ne sont pas indépendants (échantillons appariés), comparer deux moyennes revient à comparer leur différence à zéro (test sur un échantillon).

Exemple : Le taux d'albumine sérique (en g/l) des enfants hospitalisés à LWIRO est donné par le tableau ci- après :

Décédés			Vivants		
n_1	$\overline{x}_1$	s_{x_1}	n_2	$\overline{x}_2$	s_{x_2}
170	17,4	7,3	862	25,2	8,3

Le taux d'albumine sérique chez les enfants décédés est-il significativement différent de celui des enfants vivants au seuil de 5 % ? Vérifiez-le à l'aide d'un test statistique puis à l'aide d'un intervalle de confiance.

Solution

Les hypothèses du test d'égalité des variances sont :

$H_0 : \sigma_1^2 = \sigma_2^2$ (Les variances sont égales)

$H_1 : \sigma_1^2 \neq \sigma_2^2$ (Les variances sont inégales)

La statistique de test est :

$$F^{obs} = \frac{(8,3)^2}{(7,3)^2} = \frac{68,89}{53,29} = 1,292738 \simeq 1,29 < 3$$

Les variances sont donc considérées comme homogènes.

La variance poolée vaut :

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{169(7,3)^2 + 861(8,3)^2}{170+862-2} = \frac{169 \times 53,29 + 861 \times 68,89}{170+862-2} \simeq 66,3 \text{ à } 1030 \, ddl$$

$$\Rightarrow s_p = \sqrt{s_p^2} = \sqrt{66,3} = 8,14$$

Les hypothèses du test d'égalité des moyennes sont :

$H_0 : \mu_1 = \mu_2$ (Les moyennes sont égales)

$H_1 : \mu_1 \neq \mu_2$ (Les moyennes sont différentes)

En supposant que H_0 soit vraie et que les variances soient égales, la statistique de test est :

$$t^{obs} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{|17,4 - 25,2|}{8,14\sqrt{\frac{1}{170} + \frac{1}{862}}} = 11,4 \text{ à } 1030 \, ddl$$

p-value<0,001<0,05

⇒ p-value<0,05

Décision : Rejet de H0 (test significatif)

Conclusion : Au seuil de 5 %, le taux moyen d'albumine sérique chez les enfants décédés est significativement différent de (inférieur à) celui des enfants sortis vivants.

L'intervalle de confiance à 95 % de la différence vraie entre le taux moyen d'albumine sérique chez les enfants décédés et celui des enfants vivants est donné par :

$$\begin{aligned} IC_{95\%}\left(\mu_1-\mu_2\right) &= \left|\bar{x}_1-\bar{x}_2\right| \pm t_{n_1+n_2-2;0,975}\, s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \\ &= \left|17,4-25,2\right| \pm 1,96\times 8,14\sqrt{\frac{1}{170}+\frac{1}{862}} \\ &= 7,8 \pm 1,34 \\ &= \left[6,46;9,14\right] \end{aligned}$$

Au seuil de 5 %, la différence vraie entre le taux moyen d'albumine sérique chez les enfants décédés et celui des enfants sortis vivants est comprise entre 6,46 mg/dl et 9,14 mg/dl. Cet intervalle ne contient pas la valeur zéro.

Décision : Rejet de H_0 (IC significatif)

Conclusion : Au seuil de 5 %, le taux moyen d'albumine sérique chez les enfants décédés est significativement différent de (inférieur à) celui des enfants vivants.

6.6.4. Test sur deux proportions

Soient :

n_1 la taille du premier échantillon
n_2 la taille du deuxième échantillon
p_1 la proportion observée dans le premier échantillon
$q_1 = 1 - p_1$ le complément de p_1
p_2 la proportion observée dans le deuxième échantillon
$q_2 = 1 - p_2$ le complément de p_2
π_1 la proportion (vraie) dans la première population
π_2 la proportion (vraie) dans la deuxième population

Il s'agit d'un test t de Student sur deux échantillons indépendants (comparaison de deux proportions de deux échantillons indépendants).
Les hypothèses de test sont :

$H_0 : \pi_1 = \pi_2$ (Les proportions sont égales) (6.55.a)

$H_1 : \pi_1 \neq \pi_2$ (Les proportions sont inégales) (6.55.b)

En supposant que H_0 soit vraie, la statistique de test est [16] :

$$t^{obs} = \frac{|p_1 - p_2|}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n2}\right)}} \grave{a} \left(n_1 + n_2 - 2\right) ddl \tag{6.56}$$

avec

$$avec \begin{cases} \overline{p} = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2} \\ \overline{q} = 1 - \overline{p} \end{cases} \tag{6.57}$$

La p-value sera déterminée en utilisant la distribution z. La conclusion portera sur le rejet de H_0 (test significatif) si la $p-$value est inférieure ou égale au seuil de signification choisie et au non rejet de H_0 (test non significatif) si la $p-$value est supérieure au seuil de signification choisie.

L'intervalle de confiance à 95 % de la différence vraie entre les 2 proportions (si les échantillons sont indépendants) est donné par :

$$IC_{95\%}\left(\pi_1 - \pi_2\right) = |p_1 - p_2| \pm z_{1-\frac{\alpha}{2}} \sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n2}\right)} \tag{6.58}$$

Exemple

Le directeur d'un collège affirme que les élèves de son collège (groupe 1) se trouvent plus facilement un emploi d'été que les élèves du collège voisin (groupe 2). Un échantillon aléatoire de 200 élèves du groupe 1 montre que 55 d'entre eux se sont trouvé un emploi d'été alors que, pour un échantillon de 150 élèves du groupe 2, le nombre est 40. Au seuil de 1 %, le directeur a-t-il raison? Vérifiez-le à l'aide d'un test statistique, puis à l'aide d'un intervalle de confiance à 99 %. N'oubliez pas de vérifier les conditions de validité de votre intervalle de confiance.

Solution

La proportion d'élèves qui se sont trouvé un emploi d'été pour le premier groupe est :

$$p_1 = \frac{o_1}{n_1} = \frac{55}{200} = 0{,}275$$

Son complément est :

$$q_1 = 1 - p_1 = 1 - 0{,}275 = 0{,}725$$

La proportion d'élèves qui se sont trouvé un emploi d'été pour le deuxième groupe est :

$$p_2 = \frac{o_2}{n_2} = \frac{40}{150} = 0,267$$

Son complément est :

$$q_2 = 1 - p_2 = 1 - 0,267 = 0,633$$

Les hypothèses de test sont :

$H_0 : \pi_1 = \pi_2$ (Les proportions sont égales)

$H_1 : \pi_1 \neq \pi_2$ (Les proportions sont inégales)

En supposant que H_0 soit vraie, la statistique de test est :

$$t^{obs} = \frac{|p_1 - p_2|}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1} + \frac{1}{n2}\right)}} \text{ à } n_1 + n_2 - 2 \text{ ddl}$$

avec

$$avec \begin{cases} \overline{p} = \dfrac{200\dfrac{55}{200} + 150\dfrac{40}{150}}{200 + 150} = \dfrac{95}{350} = 0,257 \\ \overline{q} = 1 - 0,257 = 0,743 \end{cases}$$

soit

$$t^{obs} = \frac{|p_1 - p_2|}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1} + \frac{1}{n2}\right)}} = \frac{|0,275 - 0,267|}{\sqrt{0,257 \times 0,243\left(\frac{1}{200} + \frac{1}{150}\right)}} = \frac{0,008}{0,047} = 0,17 \text{ à } 348 \text{ ddl}$$

0,01<0,50<p-value<0,90

⇒ p-value>0,01

Décision : Non rejet de H0 (test non significatif)

Conclusion : Au seuil de 1 %, la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le premier groupe (son collège) n'est pas significativement différente de (supérieure à) la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le deuxième groupe (collège voisin).

L'intervalle de confiance à 95 % de la différence entre la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le premier groupe (son collège) et la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le deuxième groupe (collège voisin) est :

$$\begin{aligned} IC_{95\%}(\pi_1-\pi_2) &= |p_1-p_2| \pm z_{1-\frac{\alpha}{2}}\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n2}\right)} \\ &= |0,275-0,267| \pm 1,96\sqrt{0,257\times 0,243\left(\frac{1}{200}+\frac{1}{150}\right)} \\ &= 0,008 \pm 0,093 \\ &= [-0,085;0,101] \end{aligned}$$

Au seuil de 1 %, la différence entre la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le premier groupe (son collège) et la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le deuxième groupe (collège voisin) est comprise entre 0 % et 10,1 %. Cet intervalle contient la valeur zéro.

Décision : Non rejet de H0 (IC non significatif)

Conclusion : Au seuil de 1 %, la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le premier groupe (son collège) n'est pas significativement différente de (supérieure à) la proportion vraie d'élèves qui se sont trouvé un emploi d'été pour le deuxième groupe (collège voisin).

6.7. Exercices d'application

Exercice 1 :

L'an dernier, on a observé sur un échantillon de 29 appartements de 3 pièces situés en ville des dépenses de chauffage égales en moyenne à 325 dollars, avec un écart-type égal à 26 dollars. Cette année, pour un nouvel échantillon de 31 appartements de 3 pièces en ville, on a trouvé des valeurs respectives de 338 dollars et 28 dollars. N'y a-t-il pas eu une augmentation des dépenses, en moyenne, entre les deux années ? Vérifiez-le à l'aide d'un test statistique, puis à l'aide d'un intervalle de confiance à 95 % (2 méthodes).

Exercice 2 :

Une étude approfondie a évalué à 69800 euros/an le revenu moyen imposable par ménage résidant

à Neuilly-sur-Seine. Une enquête est effectuée auprès de 500 ménages pris au hasard, afin de contrôler le résultat de l'étude. Dans l'enquête, on trouve une moyenne de 68750 euros/an avec un

écart-type de 10350 euros/an. Ce résultat est-il significativement inférieur à celui trouvée par l'étude ? Vérifiez-le à l'aide d'un test statistique, puis à l'aide d'un intervalle de confiance à 95 %.

Tracez la fonction puissance de ce test. N'oubliez pas de vérifier les conditions de validité de votre intervalle de confiance.

Exercice 3 :

On administre un test à 250 psychologues et à 250 psychiatres pour mesurer leurs tendances psychopathes. Pour l'échantillon de psychologues, on obtient $x1 = 50,7$ et $1\ s = 5,2$ tandis que pour les psychiatres, $2\ x = 49,8$ et $2\ s = 5,1$. Peut-on conclure, à un seuil de signification de 5 %, que les psychologues sont plus psychopathes (i.e. ont un score plus élevé) que les psychiatres ? Vérifiez-le à l'aide d'un test statistique, puis à l'aide d'un intervalle de confiance à 95 %.

Exercice 4 :

Un étudiant voudrait savoir s'il vaut la peine de se porter candidat à la présidence de l'Association des étudiants/es de l'Université. Pour cela, un échantillon de 50 étudiants/es a montré que 22 % des étudiants/es voteraient pour lui.

a) Estimez le véritable pourcentage à un niveau de confiance de 99 %.
b) Est-il raisonnable de dire que ce pourcentage est significativement inférieur à 50 % théoriquement attendu ? Vérifiez-le à l'aide d'un test statistique.

Exercice 5 :

Des ampoules dont la durée de vie moyenne est de 750 heures sont entreposées depuis plusieurs années. On craint que ce long séjour n'ait réduit la durée de vie des ampoules. En supposant que la durée de vie se distribuée normalement, on prend 10 ampoules dont la durée de vie moyenne est 710 heures avec un écart-type de 40 heures. Au seuil de 0,10, doit-on conclure que la durée de vie des ampoules entreposées est significativement réduite ?

Annexes

Annexe 1 : Table de la loi normale standard

La table ci-dessous comporte les valeurs de la fonction de répartition de la loi normale, à savoir les valeurs de :

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	0,7703	0,7734	0,7764	0,7793	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	0,8906	0,8925	0,8943	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	0,9986	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
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Utilisation de la table de la loi normale standard

On lit les décimales dans les lignes, et les centièmes en colonnes, Par exemple, la valeur de F(1,65) se trouve à l'intersection de la ligne 1,6 et de la colonne 0,05 - on trouve F(1,65)=0,9505, à 10^{-4} près, Pour les valeurs négatives de x, on utilise la relation F(-x)=1-F(x),

Annexe 2 : Table de la loi de Student

En fonction du nombre de degrés de liberté (qu'on lit sur la première colonne) et du risque d'erreur a (qu'on lit sur la première ligne), on trouve la valeur de l'écart t qui possède la probabilité a d'être dépassé en valeur absolue,

	0,50	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001	0,0001
1	1,000	3,078	6,314	12,706	31,281	63,657	127,32	318,31	636,62	6366,2
2	0,816	1,886	2,920	4,303	6,965	9,925	14,089	22,327	34,599	99,992
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,215	12,924	28,000
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610	15,544
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,893	6,869	11,178
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959	9,082
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408	7,885
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041	7,120
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781	6,594
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587	6,211
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437	5,921
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318	5,694
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221	5,513
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140	5,363
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073	5,239
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015	5,134
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965	5,044
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922	4,966
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883	4,897
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850	4,837
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819	4,784

22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792	4,736
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768	4,693
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745	4,654
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725	4,619
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646	4,482
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591	4,389
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551	4,321
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520	4,269
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496	4,228
60	0,679	1,296	1,671	2,000	2,390	2,660	2,915	3,232	3,460	4,169
70	0,678	1,294	1,667	1,994	2,381	2,648	2,899	3,211	3,435	4,127
80	0,678	1,292	1,664	1,990	2,374	2,639	2,887	3,195	3,416	4,096
90	0,677	1,291	1,662	1,987	2,368	2,632	2,878	3,183	3,402	4,072
100	0,677	1,290	1,660	1,984	2,364	2,626	2,871	3,174	3,390	4,053
150	0,676	1,287	1,655	1,976	2,351	2,609	2,849	3,145	3,357	3,998
200	0,676	1,286	1,653	1,972	2,345	2,601	2,839	3,131	3,340	3,970
300	0,675	1,284	1,650	1,968	2,339	2,592	2,828	3,118	3,323	3,944
500	0,675	1,283	1,648	1,965	2,334	2,586	2,820	3,107	3,310	3,922
1000	0,675	1,282	1,646	1,962	2,330	2,581	2,813	3,098	3,300	3,906
infini	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291	3,891

Annexe 3 : Table de la loi du chi-carré de Pearson

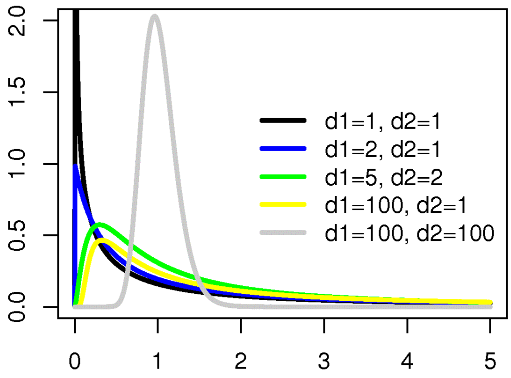
En fonction du nombre de degrés de liberté (qu'on lit sur la première colonne) et du risque d'erreur a (qu'on lit sur la première ligne), on trouve la valeur de l'écart χ^2 qui possède la probabilité a d'être dépassée,

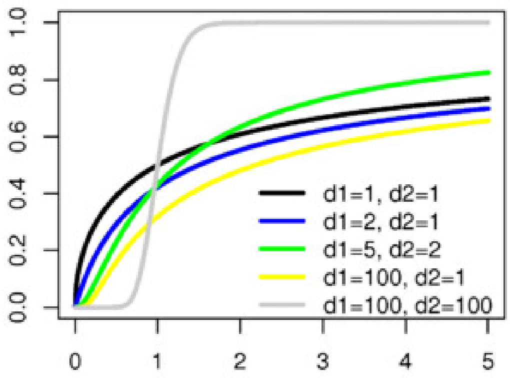
	0,10	0,05	0,01	0,001
1	2,706	3,841	6,635	10,828
2	4,605	5,991	9,210	13,816
3	6,251	7,815	11,345	16,266
4	7,779	9,488	13,277	18,467
5	9,236	11,070	15,086	20,515
6	10,645	12,592	16,812	22,458
7	12,017	14,067	18,475	24,322
8	13,362	15,507	20,090	26,124
9	14,684	16,919	21,666	27,877
10	15,987	18,307	23,209	29,588
11	17,275	19,675	24,725	31,264
12	18,549	21,026	26,217	32,909
13	19,812	22,362	27,688	34,528
14	21,064	23,685	29,141	36,123
15	22,307	24,996	30,578	37,697
16	23,542	26,296	32,000	39,252
17	24,769	27,587	33,409	40,790
18	25,989	28,869	34,805	42,312
19	27,204	30,144	36,191	43,820
20	28,412	31,410	37,566	45,315

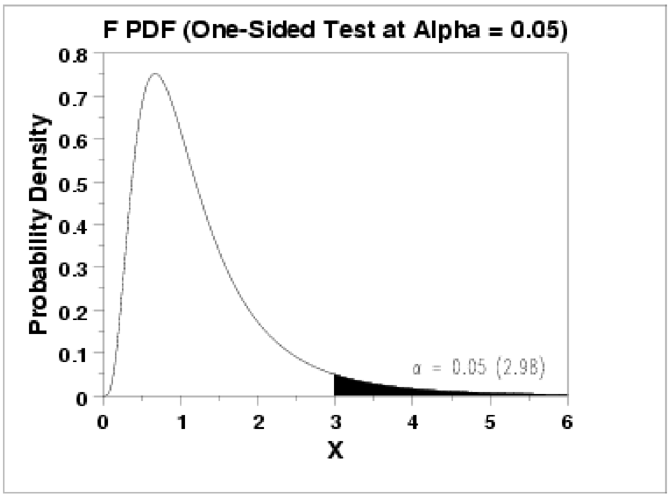
21	29,615	32,671	38,932	46,797
22	30,813	33,924	40,289	46,268
23	32,007	35,172	41,638	49,728
24	33,196	36,415	42,980	51,179
25	34,382	37,652	44,314	52,620
30	40,26	43,77	50,89	59,70
35	46,06	49,80	57,34	66,62
40	51,81	55,76	63,69	73,40
45	57,51	61,66	69,96	80,08
50	63,17	67,50	76,15	86,66
60	74,40	79,08	88,38	99,61
70	85,53	90,53	100,43	112,32
80	96,58	101,88	112,33	124,84
90	107,57	113,15	124,12	137,21
100	118,50	124,34	135,81	149,45

Annexe 4 : Table de la loi de Fisher

Densité de probabilité/ fonction de masse **Fonction de répartition**







Upper critical values of the F distribution

for ν_1 numerator degrees of freedom and ν_2 denominator degrees of freedom

5% significance level

$$F_{.05}(\nu_1, \nu_2)$$

ν_2 \ ν_1	1	2	3	4	5	6	7	8
1	161,448	199,500	215,707	224,583	230,162	233,986	236,768	238,882
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818

6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767

Références bibliographiques

1. Dagnelie P. Diversité et unité de la statistique. Journal de la société française de statistique. 1982;123(2):86‑92.

2. Saporta G. Probabilités, analyse des données et statistique. Paris: Editions Technip; 2008.

3. Dress F. Les probabilités et la statistique de A à Z: 500 définitions, formules et tests d'hypothèse. Dunod; 2007. 206 p.

4. Green TM, Hamberg CL. Pascal's Triangle, 2nd Edition. 2 edition. CreateSpace Independent Publishing Platform; 2012. 410 p.

5. Dalang RC, Conus D. Introduction à la théorie des probabilités. Lausanne: Presses polytechniques et universitaires romandes; 2015.

6. Saporta G. Probabilités analyse et des données et statistique. Paris: Technip; 2011.

7. Ross SM. Initiation aux probabilités. Presses polytechniques et universitaires romandes; 1996. 458 p.

8. Tassi P, Legait S. Théorie des probabilités en vue des applications statistiques. Editions TECHNIP; 1990. 374 p.

9. Reischer C, éditeur. Théorie des probabilités: problèmes et solutions. Sainte-Foy: Presses de l'Université du Québec; 2002. 437 p.

10. Bogaert P. Probabiblités pour scientifiques et ingénieurs: introduction au calcul des probabilités : cours et exercices corrigés. Bruxelles: De Boeck; 2006.

11. Cantoni E, Huber P, Ronchetti EM. Maîtriser l'aléatoire: exercices résolus de probabilités et statistique. Paris Berlin Heidelberg [etc.]: Springer; 2006. 256 p. (Collection Statistiques et probabilités appliquées).

12. Stone JV. Bayes' rule: a tutorial introduction to Bayesian analysis. First edition, third printing [with corrections]. Sheffield: Sebtel Press; 2014. 170 p.

13. Morris D, Bayes T. Bayes' theorem: a visual introduction for beginners. Erscheinungsort nicht ermittelbar: Blue Windmill Media; 2017. 1 p.

14. Wang H, Zheng H. Positive Predictive Value. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H, éditeurs. Encyclopedia of Systems Biology. New York, NY: Springer; 2013 [cité 7 avr 2020]. p. 1723‑4. Disponible sur: https://doi.org/10.1007/978-1-4419-9863-7_256

15. Cramér H. Random variables and probability distributions. Cambridge; New York: Cambridge University Press; 2004. 132 p.

16. Lejeune M. Statistique: la théorie et ses applications. Deuxième éd. avec exercices corrigés. Paris Berlin Heidelberg [etc.]: Springer; 2010. 446 p. (Statistique et probabilités appliquées).

17. Lecoutre J-P. Statistique et probabilités. 7e éd. Malakoff: Dunod; 2019. 317 p. (Éco Sup).

18. Petrov VV. Sums of Independent Random Variables. Berlin, Heidelberg: Springer Berlin Heidelberg; 1975 [cité 7 avr 2020]. 358 p. Disponible sur: http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=3089976

19. WELCH BL. The generalisation of student's problems when several different population variances are involved. Biometrika 1947;34(1-2):28-35. [cité 7 avr 2020]; Disponible sur: https://www.ncbi.nlm.nih.gov/pubmed/20287819

20. Newcombe RG. Confidence intervals for proportions and related measures of effect size. Boca Raton: CRC Press, Taylor & Francis Group; 2013. 442 p. (Chapman & Hall/CRC biostatistics series).

21. Srivastava MK, Srivastava N. Statistical inference: testing of hypotheses. New Delhi (India): PHI Learning; 2009. 416 p.

22. Altman DG, Machin D, Bryant TN, Gardner MJ, éditeurs. Statistics with confidence: confidence intervals and statistical guidelines. 2. ed. London: BMJ; 2000. 240 p.