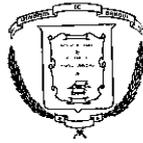


UNIVERSITE DE BANGUI

\*\*\*\*\*



FACULTE DES SCIENCES ECONOMIQUES ET  
DE GESTION

\*\*\*\*\*

DEPARTEMENT DES SCIENCES  
ECONOMIQUES

REPUBLIQUE CENTRAFRICAINE

\*\*\*\*\*

Unité – Dignité – Travail

# STATISTIQUE INFERENCELLE

*A l'usage des économistes, gestionnaires et des deuxième années des écoles de commerce*

Mexan-Ruddy-Josip ADOUM-KAMATA

UB 2018

# TABLE DES MATIERES

INTRODUCTION GENERALE.....	9
PREMIERE PARTIE: BASES PROBABILISTES .....	12
CHAPITRE I : THEORIE ET CALCUL DES PROBABILITES .....	13
I.    VOCABULAIRE DES PROBABILITES.....	13
1.1.  Epreuve ou expérience .....	13
1.2.  Ensemble fondamentale (univers) et éventualité.....	13
1.3.  Evènement :.....	13
1.3.1.  Evènement élémentaire : .....	14
1.3.2.  Evènement impossible et certain:.....	14
1.3.3.  Evènement contraire ou complémentaire. ....	14
II.   DEFINITION DE LA PROBABILITE.....	14
2.1.  Approche intuitive de la notion de probabilité .....	15
2.1.1.  Présentation .....	15
2.1.2.  Extension de la définition.....	15
2.2.  Généralisation de la notion de probabilité : approche axiomatique .....	16
2.2.1.  Algèbre d'évènement .....	16
2.2.2.  Intersection d'événements.....	16
2.2.3.  Réunion d'événements .....	16
2.2.4.  Complémentarité d'une réunion et d'une intersection .....	16
2.2.5.  Algèbre de Boole.....	17
2.2.6.  Probabilité sur un ensemble fini non vide .....	18
2.2.6.1.  Définition.....	18
2.2.6.2.  Axiomes de probabilité .....	18
2.2.6.3.  Conséquences de la définition .....	18
2.3.  Axiomes des probabilités totales .....	19
2.3.1.  Définition.....	19
2.3.2.  Généralisation.....	20
III.  AXIOME DES PROBABILITES COMPOSEES .....	20
3.1.  Probabilité conditionnelle.....	20
3.1.1.  Définition.....	20
3.1.2.  Propriétés.....	20
3.2.  Probabilités composées .....	20
3.3.  Evènements indépendants .....	21
3.3.1.  Cas de deux événements.....	21



3.3.2.	Cas de plusieurs événements .....	22
3.5.	Théorème De Bayes .....	22
CHAPITRE II : VARIABLES ALEATOIRES .....		23
I.	DEFINITION D'UNE VARIABLE ALEATOIRE .....	23
II.	VARIABLES ALEATOIRES DISCRETES .....	24
2.1.	Loi de probabilité ou fonction de distribution .....	24
2.1.1.	Définition.....	24
2.1.2.	Propriété .....	24
2.2.	Fonction de répartition .....	25
2.2.1.	Définition.....	25
2.2.2.	Propriétés.....	25
2.3.	Variables aléatoires indépendantes.....	25
2.3.1.	Définition.....	25
2.4.	Caractéristiques d'une variable aléatoire discrète .....	26
2.4.1.	Espérance mathématique .....	26
2.4.1.1.	Propriétés de l'opérateur espérance mathématique .....	26
2.4.2.	Variance et écart type .....	26
2.4.2.1.	Variance.....	26
	• Définition.....	26
	• Théorème de Koenig .....	26
	• Propriété de la variance .....	26
2.4.2.2.	Ecart type.....	27
2.4.3.	Covariance de deux variables définies sur un même univers .....	27
2.4.4.	Propriétés de la covariance .....	27
2.4.5.	Moments non centrés et centrés .....	27
III.	VARIABLES ALEATOIRES CONTINUES .....	27
3.1.	Fonction de répartition .....	27
3.1.1.	Définition.....	28
3.1.2.	Propriétés.....	28
3.2.	Loi de densité .....	28
3.2.1.	Définition : .....	28
3.2.2.	Propriétés.....	28
3.2.3.	Cas où X prend ses valeurs dans l'intervalle I de bornes $\alpha$ et $\beta$ .....	29
3.3.	Caractéristiques d'une variable aléatoire.....	29
3.3.1.	Espérance mathématique .....	29
3.3.2.	Variance.....	29

3.3.3.	Ecart type.....	29
3.3.4.	Propriétés.....	29
IV.	CONCLUSION .....	29
CHAPITRE III : LOIS USUELLES DE PROBABILITES .....		30
I.	LES LOIS FONDAMENTALES DISCRETES DE PROBABILITES .....	30
1.1.	Loi de Bernoulli.....	30
1.1.1.	Modèle de base : Epreuve de Bernoulli.....	30
1.1.2.	Définition.....	30
1.1.3.	La fonction de répartition .....	31
1.1.4.	Valeurs caractéristiques.....	31
1.2.	Loi binomiale .....	31
1.2.1.	Définition.....	31
1.2.2.	Valeurs caractéristiques.....	31
1.2.3.	Lecture de la table associée à une loi binomiale notée $Bn, p$ pour $n \leq 10$ .....	31
1.2.4.	Comportement asymptotique de la variable binomiale .....	32
1.2.5.	Somme de variables aléatoires binomiales indépendantes .....	32
1.3.	Loi hypergéométrique .....	32
1.3.1.	Définition.....	32
1.3.2.	Valeurs caractéristiques.....	33
1.3.3.	Comportement asymptotique de la variable hypergéométrique .....	33
1.4.	Loi de Poisson .....	33
1.4.1.	Génération de la loi de Poisson: Processus poissonnien .....	33
1.4.2.	Définition : .....	34
1.4.3.	Valeurs caractéristiques.....	34
1.4.4.	Somme de variables aléatoires de poisson indépendantes.....	34
1.4.5.	Lecture de la table associée à une loi de Poisson .....	34
1.4.6.	Approximation d'une loi binomiale par une loi de poisson .....	34
II.	LOIS FONDAMENTALES CONTINUES DE PROBABILITES .....	35
2.1.	Loi uniforme.....	35
2.1.1.	Définition.....	35
2.1.2.	fonction de répartition .....	35
2.1.3.	Espérance mathématique, variance et écart type .....	35
2.2.	Loi exponentielle.....	35
2.2.1.	Définition.....	35
2.2.2.	Fonction de répartition .....	35
2.2.3.	Espérance mathématique, variance et écart type .....	36



2.3.	Loi normale ou loi de Laplace-Gauss.....	36
2.3.1.	Définition de la loi normale.....	36
2.3.2.	fonction de répartition .....	36
2.3.3.	Espérance mathématique, variance et écart type .....	36
2.4.	Loi normale centrée réduite.....	37
2.4.1.	Définition.....	37
2.4.2.	Fonction de répartition .....	37
2.4.3.	Propriétés :.....	37
2.4.4.	Espérance, variance et écart type.....	38
2.4.5.	Lecture de tables de la loi normale centrée et réduite .....	38
2.4.6.	Table de la fonction de répartition de la loi normale centrée réduite .....	38
2.4.7.	Table de l'écart réduit.....	38
2.4.8.	Opérations sur les variables aléatoires normales.....	38
2.4.8.1.	Fonction affine d'une variable normale.....	38
2.4.8.2.	Somme et différence de deux variables aléatoires normales indépendantes .....	38
2.4.9.	Le théorème central limite.....	38
2.4.10.	Approximations par une loi normale.....	39
2.4.10.1.	Approximation d'une loi binomiale par une loi normale .....	39
2.4.10.2.	Approximation d'une loi de Poisson par une loi normale .....	39
III.	LOIS DE PROBABILITES DERIVEES .....	39
3.1.	Loi Gamma.....	39
3.1.1.	Définition.....	39
3.1.2.	Propriété de $\Gamma p$ .....	39
3.1.3.	Valeurs caractéristiques.....	40
3.1.4.	Somme de deux variables aléatoires indépendantes.....	40
3.2.	Loi du khi – deux $\chi^2_n$ .....	40
3.2.1.	Définition.....	40
3.2.2.	Forme de la distribution .....	40
3.2.3.	Valeurs caractéristiques.....	41
	– Espérance mathématique .....	41
	– <i>Variance</i> .....	41
3.2.4.	Somme de deux variables indépendantes de loi de Khi- Deux .....	41
3.2.5.	Table de la loi de khi-deux .....	41
3.2.6.	Approximation de la loi de khi-deux par la loi normale.....	41
3.3.	Loi de Student .....	41
3.3.1.	Définition.....	41

3.3.2.	Valeurs caractéristiques .....	42
-	Espérance mathématique .....	42
-	Variance .....	42
3.3.3.	Comportement asymptotique.....	42
4.4.5.	Table de loi de Student .....	42
4.5.	Loi de Fischer – Snedecor .....	42
4.5.1.	Définition.....	42
4.5.1.	Espérance et variance .....	42
4.5.2.	Table de la loi de Fisher – Snedecor.....	42
SECONDE PARTIE : CONCEPTS ET PROCEDURES DE L'INFERENCE STATISTIQUE .....		43
CHAPITRE IV : THEORIE D'ECHANTILLONNAGE.....		44
I.	VOCABULAIRE .....	44
II.	METHODES D'ECHANTILLONNAGE.....	45
2.1.	Méthodes d'échantillonnage probabiliste.....	45
2.1.1.	Echantillonnage Aléatoire Simple (EAS).....	45
2.1.2.	Echantillonnage aléatoire stratifié .....	46
2.1.2.1.	Répartition proportionnelle .....	46
2.1.2.2.	Répartition optimale .....	47
2.1.3.	Echantillonnage par grappes.....	48
2.1.4.	Echantillonnage systématique .....	48
2.2.	Méthodes d'échantillonnage non-probabiliste.....	49
2.2.1.	Echantillonnage de commodité (ou convenance).....	49
2.2.2.	Echantillonnages subjectifs .....	49
III.	DISTRIBUTION D'ECHANTILLONNAGES .....	50
3.1.	Position du problème.....	50
3.2.	Exemple introductif.....	50
3.3.	Distribution d'échantillonnage .....	51
3.3.1.	Notation .....	51
3.3.2.	Illustration graphique .....	52
3.3.3.	Interprétation probabiliste .....	52
3.3.4.	distribution d'échantillonnage de $X$ .....	52
•	Esperance mathématique de $X$ .....	52
•	Ecart-type .....	53
•	Pour une population finie .....	53
•	Pour une population infinie.....	53
•	Remarque : .....	53

3.3.5.	Forme de la distribution d'échantillonnage de $x$ .....	54
•	Théorème central limite.....	54
3.3.6.	distribution d'échantillonnage de $p$ .....	54
a)	Esperance mathématique de $p$ .....	54
b)	Ecart-type de $p$ .....	54
•	Pour une population finie .....	54
•	Pour une population infinie .....	54
3.3.7.	Application du théorème central limite .....	55
3.3.8.	Distribution spécifiée .....	55
CHAPITRE V : ESTIMATION .....		56
I.	ESTIMATION PONCTUELLE D'UN PARAMETRE .....	56
1.1.	Définition.....	56
1.2.	Propriétés d'un estimateur .....	57
1.2.1.	Biais d'un Estimateur .....	57
a)	Estimateur sans biais .....	57
b)	Estimateur asymptotiquement sans biais .....	57
1.2.2.	Convergence d'un estimateur .....	57
1.2.3.	Estimateur optimal .....	58
1.2.3.1.	mesure de la précision d'un estimateur : l'Erreur Quadratique Moyenne.....	58
1.2.3.2.	Estimateur à variance minimal .....	58
1.3.	Inégalité de Cramer-Rao.....	58
1.3.1.	Définition.....	58
1.3.2.	Information de Fisher d'un échantillon ( $\theta$ réel).....	59
1.3.3	Propriétés de l'information de Fisher : .....	59
1.3.4.	Théorème :.....	59
1.4.	Estimateur efficace .....	59
II.	METHODES DE CONSTRUCTION D'UN ESTIMATEUR.....	60
2.1.	Méthode du Maximum de Vraisemblance (MMV).....	60
2.1.1.	Définition.....	60
2.1.2.	Mise en œuvre .....	61
a)	L'Approche standard .....	61
b)	la méthode <i>log-vraisemblance concentré</i> .....	62
2.2.	Méthode des Moments (MM).....	62
2.2.1.	Moments centré et moments non centrés .....	62
2.2.2.	Conditions d'application de la méthode des moments .....	63
2.2.3.	Mise en œuvre .....	63

2.2.5.	Propriétés.....	64
2.3.	Estimation ponctuelle d'une moyenne .....	64
2.4.	Estimation ponctuelle de la proportion (d'un pourcentage) .....	64
III.	ESTIMATION PAR INTERVALLE DE CONFIANCE.....	65
3.1.	Définition.....	65
3.2.	Principe de construction .....	65
3.2.1.	Intervalle unilatéral $\alpha_1 \alpha_2 = 0$ .....	65
3.2.2.	Intervalle bilatéral $\alpha_1 \alpha_2 > 0$ .....	65
3.3.	Estimation par IC d'une proportion.....	66
3.3.1.	Détermination de la taille d'échantillon n d'une IC d'une proportion. ....	69
3.4.	Intervalles associés aux paramètres de la loi normale .....	70
3.4.1.	Intervalles pour la moyenne d'une loi normale d'écart type connu. ....	70
3.4.2.	Intervalles pour la moyenne d'une loi normale d'écart type inconnu. ....	71
3.4.3.	Détermination de la taille de l'échantillon pour une population qui ne suit une pas une distribution normale.....	72
3.4.4.	Intervalles pour la variance d'une loi normale d'espérance connue.....	73
3.4.5.	Intervalles pour la variance d'une loi normale d'espérance inconnue. ....	73
VI.	CONCLUSION .....	74
CHAPITRE VI :	TEST D'HYPOTHESES.....	75
I.	GENERALITES SUR LES TESTS STATISTIQUES.....	75
1.1.	Principe d'un test statistique.....	75
1.1.1.	Méthodologie.....	75
1.1.2.	Définir les hypothèses nulle et alternative.....	76
a)	Test unilatéral .....	76
•	Test unilatéral à gauche .....	76
•	Test unilatéral à droite .....	77
b)	Test bilatéral : Tester des hypothèses dans un contexte de prise de décision.....	77
1.2.	Choix d'un test statistique .....	78
1.3.	Choix de la région critique .....	78
1.4.	Règle de décision.....	79
•	Règle de décision 1 : .....	79
•	Règle de décision 2 : .....	79
1.5.	Risque d'erreur, puissance et robustesse d'un test .....	79
1.5.1.	Risque d'erreur de première espèce .....	80
1.5.2.	Risque d'erreur de deuxième espèce .....	80
1.5.3.	Puissance et robustesse d'un test $(1 - \beta)$ .....	81



BIBLIOGRAPHIE SELECTIVE .....	82
ANNEXES .....	83

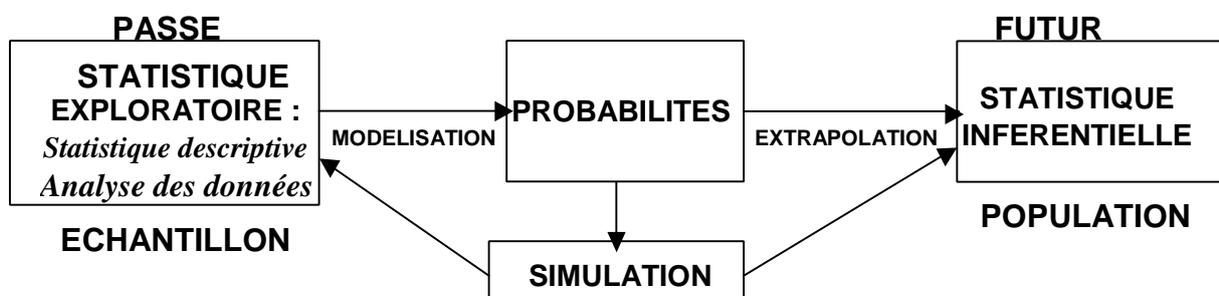
## INTRODUCTION GÉNÉRALE

Le terme statistique inférentielle d'abord désignée comme "*statistique mathématique*" (parce que la théorie des probabilités y a une large place) ou "*statistique inductive*" (parce que la démarche y est souvent inductive, plutôt que déductive, avec toute l'incertitude que cela sous-tend), prend son essor en Angleterre au début du XX<sup>e</sup> siècle, avec *Ronald A. Fisher* et *Karl Pearson*, pour répondre à des problèmes pratiques, en agronomie et en biologie.

Son but est d'*étendre* (inférer = tirer les conséquences), à *la population* toute entière, les propriétés constatées sur une partie restreinte de cette population (un échantillon). Il s'agit, ayant conscience des fluctuations d'échantillonnage, d'*estimer* un paramètre, ou de *tester une hypothèse* statistique, concernant la population étudiée. La statistique inférentielle a un aspect *décisionnel* et le calcul des probabilités y joue un rôle fondamental, en particulier pour calculer les risques d'erreur.

En effet la **probabilité**, théorie mathématique abstraite *modélisant* des phénomènes où le "hasard" intervient et la **simulation**, méthode statistique permettant la *reconstitution* fictive de l'évolution d'un phénomène sont deux domaines dont la statistique inférentielle tire sa source. En produisant des données, sous un certain modèle (probabiliste), la simulation permet d'en examiner les conséquences, soit pour juger de l'adéquation du modèle à la réalité, soit pour obtenir (ou conjecturer) des résultats difficiles, ou impossibles, à calculer. De ce fait les notions de *statistique exploratoire, probabilités, statistique inférentielle et simulation* sont souvent entre liées comme la figure ci-après nous le présente

**Figure 1** : liens entre *statistique exploratoire, probabilités, statistique inférentielle et simulation*



La *statistique exploratoire* quand elle repose sur l'*observation* de phénomènes concrets (donc *passés*). Son but est de *résumer, structurer* et *représenter (statistique descriptive)* l'information contenue dans les données. L'*analyse des données* par contre regroupe les techniques de visualisation de données multidimensionnelles (analyse en composantes principales...).

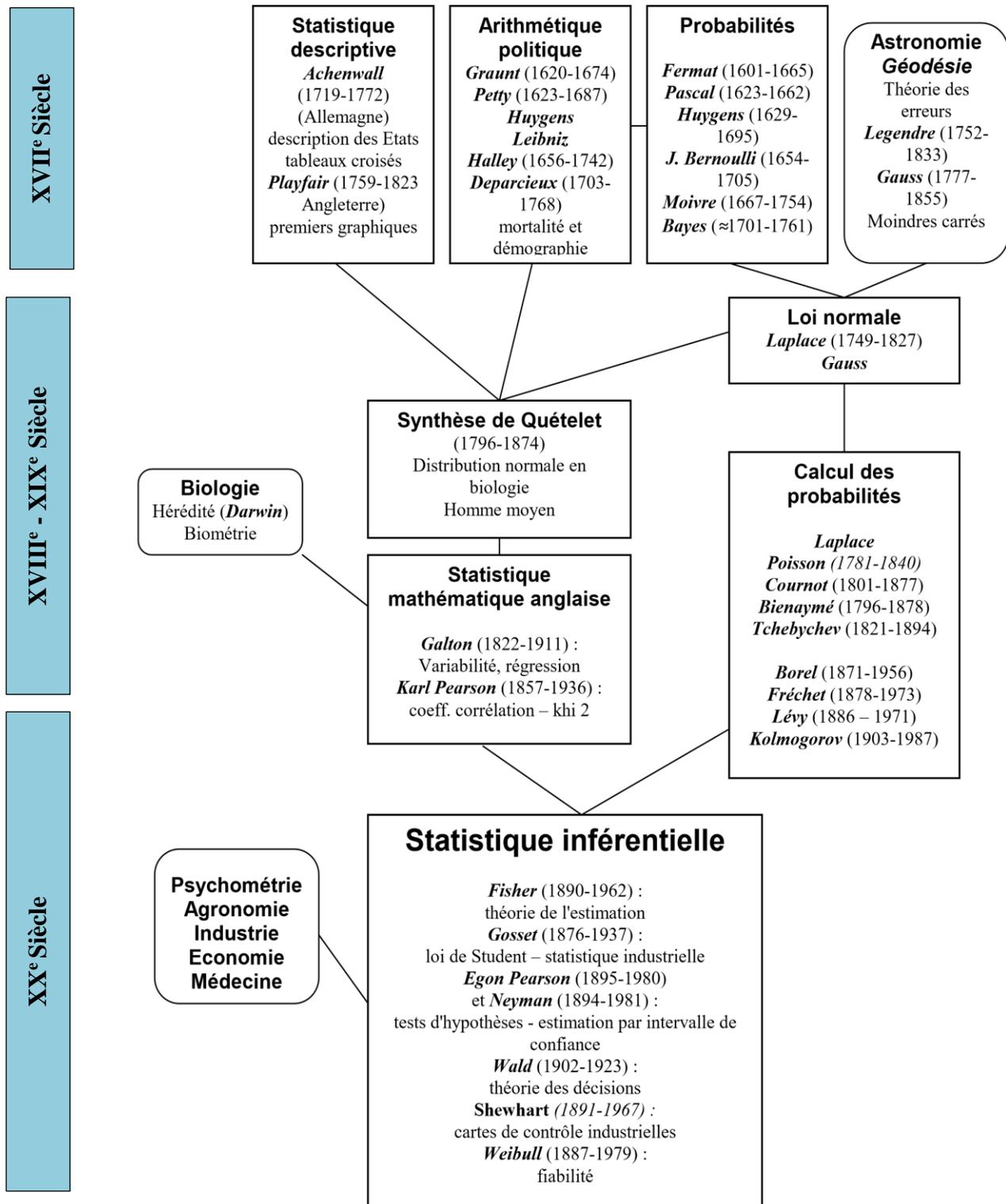
Inférer conduit donc à prendre un risque celui de généraliser à tort ; il doit être mesuré et pris en toute connaissance de cause. L'exemple classique est celui fourni par les instituts de sondage qui, un soir des élections, veulent offrir au public un résultat de portée générale, alors que le scrutin n'est que partiellement dépouillé.

Pour abordée de manière cohérente les principes de la statistique inférentielle, ce cours est subdivisé en deux parties : la première partie est destinée à introduire les bases probabilistes nécessaires à une compréhension minimale des démarches de statistiques inférentielle. Le

premier chapitre introduit les principaux concepts probabilistes, le deuxième et le troisième traitent respectivement des variables aléatoires et les lois usuelles de probabilités.

La deuxième partie quant à elle est consacrée aux principaux concepts et approches de l'inférence statistique et comporte trois chapitres. Le premier chapitre introduit la théorie d'échantillonnage, le deuxième traite de l'estimation, le troisième introduit les tests d'hypothèse.

Figure 2 : ARBRE GENEALOGIQUE DE LA STATISTIQUE INFÉRENTIELLE



# P

## REMIERE PARTIE

## BASES PROBABILISTES

La plupart des phénomènes statistiques peuvent être décrits par un petit nombre de modèles probabilistes ou lois de probabilité. Lorsque cette représentation est possible, elle fournit une description beaucoup plus riche du phénomène que le simple calcul des caractéristiques de tendance centrale et de dispersion. Elle permet notamment de calculer la probabilité de certains événements et de préciser dans une certaine mesure la représentation que l'on peut se faire de l'avenir. Il convient donc de connaître les modèles probabilistes les plus courants de façon à pouvoir rechercher celui qui est susceptible de convenir à la description d'un phénomène aléatoire déterminé. Le processus est le suivant :

- L'observation du phénomène fournit une distribution expérimentale ou empirique.
- L'analyse de cette distribution empirique (examen de représentation graphique et calcul des caractéristiques de tendance centrale et de dispersion) donne une première idée de la nature du phénomène observé. Au vu de ces premières conclusions, on choisit parmi les différents types de lois de distribution théorique celui qui paraît convenir. Cela revient à choisir la forme du « moule » dans lequel on peut « couler » le phénomène. Il faut alors, au moyen de la série empirique, estimer les paramètres de cette loi. Cela revient à choisir le « moule » de la taille qui convient.
- La substitution de la loi théorique à la distribution empirique n'est évidemment valable que si les valeurs observées et les valeurs théoriques résultant du modèle sont assez proches les unes des autres : il faut vérifier que la description donnée du phénomène par la loi théorique est acceptable, autrement dit, que les écarts observés entre les fréquences empiriques et les fréquences théoriques peuvent être raisonnablement attribués au hasard

Il nous semble donc que l'enseignement de la statistique doit respecter cette progression en trois étapes. Si l'utilité de la statistique descriptive comme première étape dans l'étude de la statistique mathématique est parfois à tort mise en doute, personne bien sûr ne conteste la nécessité d'étudier le calcul de probabilités comme deuxième étape avant d'aborder la statistique mathématique en dernière étape.

# C HAPITRE I :

## THEORIE ET CALCUL DES PROBABILITES

La théorie des probabilités est une science mathématique étudiant les lois régissant les phénomènes aléatoires. Après avoir introduit les concepts de base de l'approche probabiliste, nous présenterons des résultats fondamentaux connus sous le nom de théorème de probabilité.

### I. VOCABULAIRE DES PROBABILITES

Epreuve (ou expérience), univers, événement sont les principaux termes de base utilisés en probabilité.

#### 1.1. Epreuve ou expérience

En probabilité une Epreuve ou expérience est un processus qui génère un ensemble de résultat prédéfinis. En d'autres termes on appelle épreuve, l'action dont le résultat est soumis au hasard. Lorsque l'expérience n'est pas répétée, un seul des résultats possibles de l'expérience se produit.

**Exemple 1 :** Plusieurs exemples d'expériences, et leurs résultats possibles sont présentés ci-dessus :

Expérience	Résultats de l'expérience
Lancer une pièce de monnaie	Pile, Face
Sélectionner une pièce pour l'inspecter	Défectueuse, non défectueuse
Faire une offre de vente	Achat, pas d'achat
Lancer un dé	1, 2, 3, 4, 5, 6
Jouer au foot	Gagner, perdre, match nul

#### 1.2. Ensemble fondamentale (univers) et éventualité

L'ensemble des *résultats possibles* d'une expérience est appelée « **univers** » ou « **ensemble fondamentale** » que l'on note  $\Omega$ . Chacun de ces résultats possible est appelé une « **éventualité** ».

**Exemple 2:** Considérons la quatrième expérience inscrite dans le tableau précédent, où on lance un dé et l'on note les chiffres obtenus. L'univers associé à cette épreuve est :

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

L'élément 3 par exemple est une éventualité.

#### 1.3. Evènement :

Considérons l'ensemble fondamentale  $\Omega$  d'une expérience aléatoire. On appelle évènement, toute partie ou un sous ensemble quelconque  $A_i$  de l'univers  $\Omega$  c'est-à-dire une partie de  $\Omega$ . Un évènement peut être défini par une phrase.

**Exemple 3 :** On lance un dé et l'on note les chiffres obtenus. L'univers associé à cette épreuve est :  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . Les sous-ensembles de cet univers par exemple sont :  $A_1 = \{2, 4, 6\}$ ,  $A_2 = \{1, 2, 3, 4\}$ ,  $A_3 = \{5\}$ ,  $A_4 = \{3, 6\}$

L'événement  $A_1$  est l'événement « obtenir un chiffre paire »

### 1.3.1. Evènement élémentaire :

Chaque éventualité de l'ensemble fondamentale est un sous ensemble particulier, constitué d'un seul élément. C'est pourquoi on appelle *évènement élémentaire* noté  $\omega$  tout singleton de  $\Omega$

**Exemple 4 :** Obtenir 5 lors d'un jet de dé

$\omega_1 = \{1\}$ ,  $\omega_2 = \{2\}$ ,  $\omega_3 = \{3\}$ ,  $\omega_4 = \{4\}$ ,  $\omega_5 = \{5\}$ ,  $\omega_6 = \{6\}$ , sont des évènements élémentaires.

**NB :** un évènement est donc la réunion d'un certain nombre d'évènements élémentaires.

### 1.3.2. Evènement impossible et certain:

Parmi les sous ensemble de  $\Omega$  figurent l'ensemble vide  $\emptyset$  et  $\Omega$  lui-même. Ces deux sous-ensembles sont des évènements. Or quel que soit le résultat  $\omega$  lors d'une épreuve, il va de soi que :

- $\omega \notin \emptyset$ , donc l'évènement  $\emptyset$  n'est jamais réalisé. On l'appelle *évènement impossible*.

**Exemple 5 :** Obtenir un nombre supérieur à 7 lors d'un lancer d'un dé

- $\omega \in \Omega$  (car l'univers contient toutes les éventualités), donc l'évènement  $\Omega$  est toujours réalisé. On l'appelle *évènement certain* c'est à dire un évènement qui est réalisé pour tous les éléments de  $\Omega$

**Exemple 6 :** Dans le lancer d'un dé, l'évènement « obtenir un nombre inférieur à 7 » est un évènement certain

On note donc :

- $\emptyset$  : Evènement impossible
- $\Omega$  : Evènement certain

### 1.3.3. Evènement contraire ou complémentaire.

Soit  $A$  un évènement. L'évènement contraire noté  $\bar{A}$ , est l'évènement qui est réalisé si et seulement si  $A$  ne l'est pas.  $\bar{A}$  est donc le complémentaire dans  $\Omega$  de  $A$ .

**Exemple 7 :** Considérons les quatre évènements suivant :

$$A_1 = \{2, 4, 6\}, A_2 = \{1, 2, 3, 4\}, A_3 = \{5\}, A_4 = \{3, 6\}$$

Les évènements contraires respectifs sont :  $\bar{A}_1 = \{1, 3, 5\}$ ,  $\bar{A}_2 = \{5, 6\}$ ,  $\bar{A}_3 = \{1, 2, 3, 4, 6\}$ ,  $\bar{A}_4 = \{1, 2, 4, 5\}$

**Exemple 8 :** On dispose de 10 livres dont 3 livres de statistique et les autres de macroéconomie.

On choisit au hasard quatre de ces livres et on note :

A : l'évènement « Obtenir au moins deux livres en statistique »

B : l'évènement « Obtenir exactement deux livres de statistiques »

Déterminer les évènements contraires.

## II. DEFINITION DE LA PROBABILITE

Soient  $\Omega$  l'univers associé à une épreuve, et  $P(\Omega)$  ensemble des parties de  $\Omega$ . On appelle espace probabilisable le couple  $(\Omega, P(\Omega))$

## 2.1. Approche intuitive de la notion de probabilité

Intuitivement, la notion de la probabilité est connue par tous. Ce concept fait généralement référence à l'idée de la chance empruntée au jeu d'hasard.

### 2.1.1. Présentation

Lorsqu'on lance une pièce de monnaie, le jet peut amener soit « pile » (P), soit « face » (F). L'ensemble des éventualités, c'est-à-dire des résultats possibles, appelé ensemble fondamental ou encore univers est constitué de deux éléments :  $\Omega = \{P, F\}$

On ne peut prévoir à l'avance quel résultat sera amené. Si la pièce est *symétrique* et si elle est lancée au *hasard* c'est-à-dire loyalement, il n'y a pas de raison que le résultat P ait plus de chance d'apparaître que le résultat F. dans ces conditions, on dit que les deux éventualités :

- Apparition du résultat P
- Apparition du résultat F

Sont *également probables* ou *équiprobables* et on attribue à chacune de ces éventualités la probabilité  $\frac{1}{2}$ .

On peut rapporter cet exemple on tirage d'une carte (par exemple l'as de trèfle) parmi un jeu de 52 cartes.

D'une façon générale si les différents résultats possibles constituant l'ensemble fondamental (pile ou face, chacune des 52 cartes du jeu) qui peuvent apparaître à la suite d'une *épreuve* (jet de dé, tirage d'une carte) effectué dans un cadre *d'hypothèses* bien précisé (la pièce est symétrique et lancé loyalement, la carte est tirée loyalement d'un jeu convenablement battu) sont *également probables*, on attribue à chacune des éventualités correspondantes la probabilité  $1/N$ , égale à l'inverse du nombre N d'éléments de l'ensemble fondamental.

Nous introduisons ainsi la notion de probabilité dans le cas particulier privilégié ou à la fois :

- L'ensemble fondamental défini par l'épreuve effectuée comporte un nombre *fini* d'éléments ;
- Le cadre d'hypothèses conduit à affecter la *même* probabilité à chacun des éléments de l'ensemble fondamental pour des raisons de *symétrie* ;

C'est-à-dire où l'ensemble fondamental est constitué d'un nombre *fini* d'éventualités *équiprobables*. Dans ce cas particulier privilégié, la notion de probabilité est introduite par la définition de la *valeur numérique* attachée à chaque élément de l'ensemble fondamental et qu'on appelle sa probabilité.

### 2.1.2. Extension de la définition

Soit l'ensemble fondamental E, fini avec N éléments équiprobables. Soit A un événement quelconque de E, la probabilité d'un événement A est par définition égale au rapport du nombre  $N_A$  d'événements élémentaires équiprobables définissant A au nombre total N d'événements élémentaires équiprobables possibles :

$$P(A) = \frac{\text{nombre d'événements équiprobables définissant A}}{\text{nombre total d'événements équiprobables possibles}} = \frac{N_A}{N}$$

De manière plus simple :

$$P(A) = \frac{\text{nombre de cas favorable}}{\text{nombre de cas possibles}} = \frac{\text{card}(A)}{\text{card}(E)}$$

**Exemple 9 :** On lance deux dés non pipé et on note la somme de chiffres obtenus. Quelle est la probabilité d'obtenir 7

On tire simultanément et au hasard 5 cartes dans un jeu de 32 cartes. Quelle est la probabilité de tirer le roi de cœur.

## 2.2. Généralisation de la notion de probabilité : approche axiomatique

Historiquement, la notion de probabilité a été d'abord introduite dans le cas où on pouvait définir un système d'événements équiprobables. Mais progressivement il a fallu l'étendre pour pouvoir décrire des épreuves plus complexes où aucune considération de symétrie ne permettait d'affecter des valeurs numériques au degré de vraisemblance des divers résultats possibles. Ainsi, vers 1930, KOLMOGOROV a construit une théorie mathématique des probabilités. Basées sur trois axiomes fondamentaux, cette théorie confirme et généralise les propriétés établies à partir d'expériences. De là, le calcul des probabilités s'est avéré nécessaire dans un grand nombre de domaines (Economie, physique, médecine etc.) où la répétition à l'identique d'une épreuve est souvent impossible.

### 2.2.1. Algèbre d'évènement

Si on veut formaliser un problème dans lequel le hasard intervient, on doit construire un modèle probabiliste, choisi en fonction du but que l'on poursuit. Ce modèle est constitué d'un ensemble fondamental, d'une tribu d'évènement et d'une probabilité. Nous présentons comment certaines notions des théories d'ensembles se traduisent en termes d'évènement.

### 2.2.2. Intersection d'événements

Soit A et B deux ensembles. L'intersection de A et B noté  $A \cap B$  est l'ensemble des éléments de  $\Omega$  appartenant à la fois à A et B.

$A \cap B$  est l'évènement qui est réalisé que si A et B le sont

Lorsque  $A \cap B$  est vide, cela signifie que A et B ne peuvent être réalisés simultanément. On dit alors que A et B sont des événements incompatibles.

A et B sont incompatibles si et seulement si  $A \cap B \neq \emptyset$

### 2.2.3. Réunion d'événements

A et B étant deux événements, la réunion de A et B, noté  $A \cup B$ , est l'ensemble des éléments de  $\Omega$  appartenant soit à A soit à B, sans exclure les éléments appartenant à la fois à A et à B s'il y en a.

$A \cup B$  est l'évènement qui est réalisé que si A ou B l'est (le « ou » étant non exclusif)

### 2.2.4. Complémentarité d'une réunion et d'une intersection

On pourra vérifier sur un schéma les propriétés suivantes qu'il est bon de connaître

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

Et l'on peut généraliser sous la forme :

$$\overline{A_1 \cap A_2 \cap \dots \cap A_n} = \overline{A_1} \cup \overline{A_2} \cup \dots \cup \overline{A_n}$$

et

$$\overline{A_1 \cup A_2 \cup \dots \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n}$$

**Exemple 10:** on lance trois dés et on désigne par A, B, C les événements :

A : sur les 3 faces, il y a au moins un as,

B : sur les 3 faces, deux au moins sont identiques,

C : la somme des points est paire.

Evaluer les probabilités des évènements  $A, B, C, A \cap B, A \cap C, B \cap C, A \cap B \cap C$

**2.2.5. Algèbre de Boole**

La notion d'algèbre de Boole est introduite en calcul des probabilités en raison de la règle d'addition des probabilités comme on le verra on le verra plus loin. Considérons une famille  $\mathcal{F}$  de partie de l'ensemble fondamental E et envisageons les propriétés suivantes :

1. Si une partie A de E appartient à  $\mathcal{F}$ , la partie *complémentaire*  $\bar{A}$  de A par rapport à E appartient aussi  $\mathcal{F}$  :

$$(A \in \mathcal{F}) \Leftrightarrow (\bar{A} \in \mathcal{F})$$

Si cette propriété est satisfaite, on dit que  $\mathcal{F}$  est fermée par rapport à la *complémentation* relativement à E.

2. Si  $A_1, \dots, A_n$  est une suite *finie* de partie de E appartenant à  $\mathcal{F}$ , alors la *réunion* de ces parties appartient aussi a  $\mathcal{F}$ :

$$(A_1 \in \mathcal{F}, \dots, A_n \in \mathcal{F}) \Rightarrow [(A_1 \cup A_2 \cup \dots \cup A_n) \in \mathcal{F}]$$

Si cette propriété est satisfaite, on dit que  $\mathcal{F}$  est *fermée* par rapport à la *réunion* portant sur une *finité* d'éléments.

3. Si  $A_1, \dots, A_n, \dots$  est une suite infinie dénombrable de partie de E appartenant à  $\mathcal{F}$ , alors la réunion infinie de ces parties appartient aussi a  $\mathcal{F}$ :

$$(A_1 \in \mathcal{F}, \dots, A_n \in \mathcal{F}, \dots) \Rightarrow ((A_1 \cup \dots \cup A_n \cup \dots) \in \mathcal{F})$$

Si cette propriété est satisfaite, on dit que  $\mathcal{F}$  est *fermée* par rapport à la *réunion* portant sur une *infinité dénombrable* d'éléments.

Si  $\mathcal{F}$  possède les propriétés 1 et 2, on dit que  $\mathcal{F}$  est une algèbre de Boole ; Si  $\mathcal{F}$  possède les propriétés 1 et 3, on dit que  $\mathcal{F}$  est une  $\sigma$  – *algèbre* ou encore une famille de Borel.

Une  $\sigma$  – *algèbre* est une algèbre de Boole puisque la propriété 3 implique la propriété 2

**Exemple 11:** soit E un ensemble de cinq éléments notés a, b, c, d, e. une partie de E est désignée par ses éléments écrites entre parenthèses. Ainsi (a, d) désigne la partie constitué des deux éléments a et d.

On considère les deux familles  $\mathcal{F}_1$  et  $\mathcal{F}_2$  de parties de E :

$$\mathcal{F}_1: \{\emptyset, (a), (b, c, d, e), (a, b, c, d, e)\}$$

$$\mathcal{F}_2: \{\emptyset, (a), (b), (a, b), (c, d, e), (a, c, d, e), (b, c, d, e), (a, b, c, d, e)\}.$$

- a) Vérifier que  $\mathcal{F}_1$  et  $\mathcal{F}_2$  sont deux algèbre de Boole.
- b) Construire la plus petite algèbre de Boole  $\mathcal{F}_3$  contenant les parties de (a) et (c, d)
- c) Vérifier que l'intersection  $\mathcal{F}_2$  et  $\mathcal{F}_3$  est une algèbre de Boole.



Le tableau ci-après nous donne un résumé des principaux termes du vocabulaire des probabilités et leur notation (signification abrégé) :

Vocabulaire des événements	Signification des ensembles	Notation
Univers	Ensemble $\Omega$	$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$
Eventualité	Elément de $\Omega$	$\omega_i (\omega_i \in \Omega)$
Événement	Partie de $\Omega$	$E (E \subset \Omega)$
Événement élémentaire	Singleton	$\{\omega_i\} (\omega_i \in \Omega)$
Événement certain	Partie pleine	$\Omega, [P(\Omega) = 1]$
Événement impossible	Partie vide	$\phi, [P(\phi) = 0]$
Événement « A ou B »	Réunion des parties A et B	$A \cup B$
Système complet d'événements	Réunion des parties de $\Omega$	$A_1 \cup A_2 \dots \cup A_k = \Omega$
Événement « A et B »	Intersection des parties A et B	$A \cap B$
Événement A et B incompatible	Parties A et B disjointes	$A \cap B \neq \emptyset$
Événement contraire A	Complémentaire de A dans $\Omega$	$\bar{A} = \Omega - A [1 - P(A)]$
Système complet d'événements	Réunion des parties de $\Omega$	$A_1 \cup A_2 \dots \cup A_k = \Omega$

**Exemple 12:** on lance un dé et on observe le numéro tiré. On considère les événements suivants :

A : « Obtenir un nombre pair »

B : « Obtenir un nombre premier »

C : « obtenir 6 »

Déterminer :  $A \cup B, A \cap B$  puis  $B \cap C$

## 2.2.6. Probabilité sur un ensemble fini non vide

### 2.2.6.1. Définition

Soit  $\mathcal{F}$  une famille de parties de E constituant une  $\sigma$  – algèbre. Cette  $\sigma$  – algèbre est probabilisée lorsqu'on définit pour toute partie A de E appartenant à  $\mathcal{F}$  une application P dans  $\mathbb{R}$  :

$$P: \mathcal{F} \rightarrow \mathbb{R}$$

$$A \mapsto P(A)$$

Dont le résultat est appelée probabilité de A et qui satisfait aux trois conditions fondamentales :

### 2.2.6.2. Axiomes de probabilité

**Axiome 1:** La probabilité associée à l'ensemble des événements est égale à l'unité  $P(E) = 1$

**Axiome 2:** La probabilité associée à l'événement A est un nombre positif ou nul. Pour tout A:

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

**Axiome 3 :** Pour tous événements A et B incompatible, la probabilité de la réunion de ces événements est égale à la somme des probabilités de A et de B

$$P(A \cup B) = P(A) + P(B)$$

### 2.2.6.3. Conséquences de la définition

De ces axiomes, on peut déduire les propriétés suivantes :

• **Propriétés**

P étant une probabilité définie sur  $\Omega$ , on a les propriétés suivantes :

- i. La probabilité de l'événement impossible est nulle :  $P(\emptyset) = 0$
- ii.  $\bar{A}$  étant l'événement contraire de A, la probabilité de réaliser  $\bar{A}$  vaut :  $P(\bar{A}) = 1 - P(A)$
- iii. Si  $A \subset B$ , alors :  $P(A) \leq P(B)$
- iv. **Extension de l'axiome 3 :**

Si  $E_1, E_2, E_3, \dots, E_k$  sont  $k$  parties de  $E$  deux à deux disjoints supposons que l'axiome 2 se générale à  $k-1$  parties de la formule :

$$P(E_1 \cup E_2 \cup \dots \cup E_{k-1} \cup E_k) = P(E_1) + P(E_2) + \dots + P(E_k)$$

v. **Application a une partition finie de E**

Si l'ensemble  $\{E_1, E_2, E_3, \dots, E_k\}$  est une partition de E, les  $k$  parties qui le constituent sont disjointes et leur union est E, et la formule ci-dessus donne :

$$P(E_1) + P(E_2) + \dots + P(E_k) = \sum_{i=1}^k P(E_i) = 1$$

**Exemple 13:** A la sortie d'une chaîne de fabrication, des boulons sont susceptibles de présenter deux défauts. Un très grand nombre d'observations a permis d'établir que :

- La proportion de boulons fabriqués ayant le défaut a est de 5% ;
- La proportion de boulons fabriqués ayant le défaut b est de 3% ;
- La proportion des boulons fabriqués ayant les deux défauts est de 1% ;

On note : A « Événement un boulon présentant le défaut a »

B « Événement un boulon présentant le défaut b »

La probabilité que A se réalise vaut donc 0,005, la probabilité que B se réalise 0,03 et la probabilité que A et B se réalisent simultanément vaut 0,01.

Calculer la probabilité qu'un boulon présente :

- a) Le défaut a ou b ;
- b) Le défaut a seulement ;
- c) Aucun défaut

De façon générale, si les différents résultats possible constituant l'ensemble (pile ou face, chacune des 52 cartes du jeu) qui peuvent apparaître à la suite d'une épreuve

**2.3. Axiomes des probabilités totales**

**2.3.1. Définition**

Si A et B sont des événements compatibles, la probabilité de la réunion de ces deux événements est égale à la somme de leur probabilité, diminuée de la probabilité de leur intersection.

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

Cette formule est connue sous le nom du *théorème des probabilités totales*. Elle contient l'axiome 3 comme cas particulier.

**Exemple 14:** soit une urne contenant des boules dont certaines sont noires et des cubes dont certains sont noirs. On en tire un objet et on suppose connues les probabilités suivantes :

- La probabilité de tirer une boule est de 0.7

- La probabilité de tirer un objet noir est de 0.3
- La probabilité de tirer une boule noire est de 0.1

Calculer les probabilités :

- La probabilité de tirer une boule ou un objet noir
- La probabilité de tirer un cube noir

### 2.3.2. Généralisation

On peut chercher une formule analogue pour exprimer la probabilité de l'union de plus de deux événements. Prenons le cas de trois événements on a :

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_2 \cap E_3) - P(E_1 \cap E_3) + P(E_1 \cap E_2 \cap E_3)$$

Quand tous les événements sont deux à deux incompatibles on trouve la formule généralisant l'axiome 3 :

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3)$$

Finalement le cas particulier de k événements incompatibles deux à deux redonne la formule :

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3 + \dots + E_k) \\ = P(E_1) + P(E_2) + P(E_3) + \dots + P(E_k) - P(E_1 \cap E_2) - P(E_1 \cap E_3) - \dots - P(E_1 \cap E_k) \\ - P(E_2 \cap E_3) - P(E_2 \cap E_4) - \dots - P(E_2 \cap E_k) - \dots - P(E_{k-1} \cap E_k) \\ + P(E_1 \cap E_2 \cap E_3 \cap \dots \cap E_k) \end{aligned}$$

## III. AXIOME DES PROBABILITES COMPOSEES

### 3.1. Probabilité conditionnelle

Soit  $\Omega$ , l'univers sur lequel on a défini une probabilité P

#### 3.1.1. Définition

Soient A et B deux événements, avec  $P(B) > 0$ , on appelle probabilité conditionnelle de A sachant B notée  $P(A/B)$  ou  $P_B(A)$  définie par :

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

**Exemple 15 :** Reprenons l'exemple 10 sur la chaîne de fabrication. Supposons qu'à la suite d'un incident technique 50% des boulons présente le défaut a, 30% le défaut b et 10% le défaut A et B. Si l'on prend un boulon au hasard, on constate qu'il présente le défaut a, qu'elle est la probabilité qu'il présente aussi le défaut B ?

#### 3.1.2. Propriétés

(a)	Pour tout A : $0 \leq P_B(A) \leq 1$
(b)	$P_B(\Omega) = 1$
(c)	Pour tous les événements $A_1$ et $A_2$ incompatibles, $P_B(A_1 \cup A_2) = P_B(A_1) + P_B(A_2)$
(d)	$P_B(\emptyset) = 0$
(e)	$P_B(\bar{A}) = 1 - P_B(A)$
(f)	$A_1 \subset A_2 \Rightarrow P_B(A_1) \leq P_B(A_2)$
(g)	Pour tous événements $A_1$ et $A_2$ $P_B(A_1 \cup A_2) = P_B(A_1) + P_B(A_2) - P_B(A_1 \cap A_2)$

### 3.2. Probabilités composées

La formule de la probabilité composée fournit la règle de calcul de la probabilité de la réalisation simultanée de deux événements. Elle découle directement de la définition des probabilités conditionnelles.

$$P(A \cap B) = P(A) \times P(B/A)$$

Ou encore :

$$P(A \cap B) = P(B) \times P(A/B)$$

Cette relation montre comment la réalisation d'un évènement (A) modifie la probabilité d'un autre (B).

**Exemple 16 :** le parc informatique de la FASEG est composé de 20 ordinateurs dont :

- 3 sont considérés comme neuf
- 9 sont considérés comme récents
- Les autres sont considérés comme anciens

Une étude statistique indique que :

- 5% des ordinateurs neufs sont défectueux
- 10% des ordinateurs récents sont défectueux
- 20% des ordinateurs anciens sont défectueux.

On choisit au hasard un ordinateur de ce parc. On note les événements suivants

N : « Ordinateur est neuf »

R : « ordinateur est R »

A : « Ordinateur est ancien »

D : « Ordinateur est défectueux

$\bar{D}$  : « L'évènement contraire de D

1. Calculer la probabilité que l'ordinateur choisi soit neuf et défectueux
2. Déterminer que la probabilité que l'ordinateur choisi soit ancien sachant qu'il est défectueux.
3. Les événements N et D sont indépendants ?

### 3.3. Évènements indépendants

Deux cas sont à distinguer :

#### 3.3.1. Cas de deux événements

Deux événements A et B sont indépendants si la réalisation de l'un n'a pas d'influence sur la réalisation de l'autre. Dans ce cas, la probabilité de réaliser A sachant que B est réalisé doit être égale à P(A) et puisque  $P(A/B) = \frac{P(A \cap B)}{P(B)}$  avec  $P(B) \neq 0$ . On obtient donc  $P(A) = \frac{P(A \cap B)}{P(B)}$ , ce qui nous conduit à la définition suivante :

A et B sont deux événements indépendants si et seulement si

$$P(A \cap B) = P(A) \times P(B)$$

**Remarque :** Ne pas confondre « événements indépendants » et « événements incompatibles ». L'étude de l'indépendance d'évènements doit être faite à partir de la définition mathématique ci-dessus et non à partir des idées intuitives que l'on pourrait avoir

**Exemple 17:** On tire au hasard 13 cartes parmi un jeu de 52 cartes. Quelles est la probabilité de n'obtenir qu'un seul as lorsqu'on sait que parmi les 13 cartes figure l'as de pique ?

- **Propriétés**

Si A et B sont deux événements indépendants, alors :

- A et  $\bar{B}$  sont indépendants
- $\bar{A}$  et B sont indépendants
- $\bar{A}$  et  $\bar{B}$  sont indépendants

### 3.3.2. Cas de plusieurs événements

Considérons une famille  $A_1, A_2, A_3, \dots, A_n$  d'événements (contenus dans un même univers)

- Les événements  $A_1, A_2, A_3, \dots, A_n$  sont **indépendants deux à deux** si l'on a : Pour tout  $A_i$  et tout  $A_j$  avec  $i \neq j$  :

$$P(A_i \cap A_j) = P(A_i) \times P(A_j)$$

- Les événements  $A_1, A_2, A_3, \dots, A_n$  sont **mutuellement indépendants** si l'on a, pour toute sous famille  $A_{i_1}, A_{i_2}, A_{i_3}, \dots, A_{i_k}$  avec  $k \geq 2$

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

### 3.4. Formule des probabilités totales

Soit  $P$  une probabilité définie sur l'univers  $\Omega$ . Considérons une famille d'événements  $A_1, A_2, A_3, \dots, A_n$  incompatibles deux à deux, tels qu'aucun d'entre eux ne soit impossible, et dont la réunion est  $\Omega$ . Une telle famille forme un système complet d'événements.

Soit  $B$  un événement donné, on a alors :

$$P(B) = P(A_1) \times P(B/A_1) + P(A_2) \times P(B/A_2) + \dots + P(A_n) \times P(B/A_n)$$

### 3.5. Théorème De Bayes

En reprenant l'exemple ci-dessus, et si l'emballage du paquet choisi au hasard est endommagé, qu'elle est la probabilité qu'elle provienne du fournisseur  $B$  ?

Sans le mentionner, nous avons utilisé dans l'exemple ci-dessus le théorème de Bayes au cas particulier. Ainsi, il se généralise sous la forme suivante :

$$P(A_i/B) = \frac{P(A_i) \times P(B/A_i)}{P(A_1) \times P(B/A_1) + P(A_2) \times P(B/A_2) + \dots + P(A_n) \times P(B/A_n)}$$

#### Exemple 18 :

Dans une population pour laquelle 1 habitant sur 100 est atteint d'une maladie génétique  $A$ , on a mis au point un test de dépistage. Le résultat du test est soit positif ( $T$ ) soit négatif ( $\bar{T}$ )

On sait que :  $P(T/A) = 0.8$  et  $P(\bar{T}/\bar{A}) = 0.9$

On soumet un patient au test. Celui-ci est positif. Quelle est la probabilité que ce patient soit atteint de la maladie  $A$  soit  $P(A/T)$  ?

#### Solution

D'après la formule de Bayes :

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

D'où

$$P(A/T) = \frac{0.01 \times 0.8}{0.8 \times 0.01 + 0.1 \times 0.99} = 0.075$$

Ainsi *avant le test*, la probabilité *d'être malade* était de  $P(A) = 0.01$  (probabilité à priori) et *après le test* la probabilité *d'être malade* est de  $P(A/T) = 0.075$  (probabilité à posteriori).

Ainsi le test apporte un supplément d'information.

# C

## HAPITRE II :

# VARIABLES ALEATOIRES

Ce chapitre est consacré à la présentation des variables aléatoires. Après une définition axiomatique, on étudie ensuite les variables aléatoires discrètes et continues. Le vocabulaire de la statistique descriptive se trouve en calcul de probabilités à propos des variables aléatoires : fonction de répartition, densité de probabilité, courbe cumulative, histogramme, diagramme en bâton ; la différence provient de ce que la fréquence a laissé place à la probabilité.

### I. DEFINITION D'UNE VARIABLE ALEATOIRE

Soit un ensemble fondamentale  $E$  dont les éléments sont désignés par  $e$  et  $\mathcal{F}$  une  $\sigma$ -algèbre de partie de  $E$ . Soit  $P$  l'application qui associe à chaque élément  $A$  de  $\mathcal{F}$  sa probabilité :

$$P: A \subset E, A \in \mathcal{F} \rightarrow \mathbb{R}$$

$$A \mapsto P(A) \in [0, 1]$$

Soit  $X$  une application qui associe à chaque élément  $e$  de  $E$  un nombre réel  $X(e)$  :

$$X: e \in A_x \rightarrow \mathbb{R}$$

$$e \mapsto X(e) \in ]+\infty, x[$$

On dit que l'application  $X$  est une variable aléatoire si, quel que soit  $x$ ,  $A_x$  appartient à  $\mathcal{F}$ .

De manière plus simple, on appelle variable aléatoire notée  $X$ , **toute application permettant d'associer au résultat d'une expérience aléatoire (épreuve) un nombre réel**. On note  $X(\Omega)$  l'ensemble de toutes les valeurs que peut prendre  $X$ . La variable aléatoire  $X$  est dite :

➤ **Discrète** lorsque l'ensemble des valeurs qu'elle peut prendre est dénombrable

**Tableau 2.1 : Exemples de variables aléatoires**

Expériences aléatoire	Variable aléatoire (X)	Valeurs que prendre la variable aléatoire
Contacter cinq clients	Nombre de clients qui passent commande	0, 1, 2, 3, 4, 5,
Inspecter un magasin de 50 radios	Nombre de radios défectueuses	0, 1, 2, ..., 49, 50
Dispenser un cours de statistique inférentielle pendant une journée	Nombre d'étudiants	0, 1, 2, ...
Vendre de l'eau fraîche	Sexe des clients	0 si le client est un homme ; 1 si le client est une femme

**Exemple 19 :** on lance deux dés. On note  $S$  l'application qui, à chaque lancer associe la somme des résultats obtenus. Déterminer  $s_i$

➤ *Continue* lorsqu'elle peut prendre toutes les valeurs d'un intervalle  $\mathbb{R}$

**Tableau 2.2.** : Exemples de variables aléatoires continues

Expériences aléatoire	Variable aléatoire (X)	Valeurs que prendre la variable aléatoire
Gérer une banque	Temps écoulé entre les arrivées des clients en minutes	$X \geq 0$
Remplir une bouteille d'Isenbeck	Volume exacte en cl (65cl avec une marge de 0.05cl)	$64,95 \text{ cl} \leq X \leq 65,05 \text{ cl}$
Produire des beignets	Température à laquelle le beignet devient doré (min 110°C ; max 150°C)	$110^\circ \text{ C} \leq X \leq 150^\circ \text{ C}$
Evaluer les étudiants d'une classe	Pourcentage des étudiants ayant la moyenne	$0 \leq X \leq 100$

**Exemple 20** : après ensachage, on pèse des paquets de sucre. On note Y l'application qui à chaque paquet, associe son poids en gramme. On a constaté que les poids varient entre 950g et 100g.

## II. VARIABLES ALEATOIRES DISCRETES

Une variable aléatoire X est dite discrète si les différentes valeurs possibles sont en nombre fini ou en infinité dénombrable.

### 2.1. Loi de probabilité ou fonction de distribution

#### 2.1.1. Définition

On appelle loi de probabilité (ou fonction de distribution) de la variable aléatoire X, toute application qui, à chaque valeur possible  $x$  d'une variable aléatoire X associe la probabilité  $P(X = x)$ .

**Remarque** : Il faut bien distinguer la signification de  $P(X = x)$  et celle de  $(X = x)$

$(X = x)$  désigne l'événement « variable aléatoire X prend la valeur  $x$  » et ;

$P(X = x)$  est la probabilité de réaliser cet événement.

#### 2.1.2. Propriété

La somme des probabilités vaut 1 :  $\sum_{x \in X(\Omega)} P(X = x) = 1$  (cf l'axiome 2 du calcul de probabilité)

**Exemple 21** : Dans une population d'étudiant en Sciences économiques, 10% d'entre eux sont abonnés à 3 hebdomadaires, 20% à 2 hebdomadaires, 40% à un seul hebdomadaire, les autres n'ayant aucun abonnement. Déterminer les valeurs de X et déterminer sa loi de probabilité.

Les valeurs possible  $x$  de X sont  $\{0, 1, 2, 3\}$  et

$$P(X = 1) = 0,4 ; P(X = 2) = 0,2 ; P(X = 3) = 0,1$$

On sait que :

$$\sum_{x \in X(\Omega)} P(X = x) = 1 = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X = 0) = 1 - P(X = 1) - P(X = 2) - P(X = 3)$$

$$P(X = 0) = 1 - 0,4 - 0,2 - 0,1 = 0,3$$

$$P(X = 0) = 0,3$$

La loi de probabilité de X est consignée dans le tableau suivant :

$x$	0	1	2	3
$P(X = x)$	0,3	0,4	0,2	0,1

De plus  $\sum P(X = x) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1$

**2.2. Fonction de répartition**

**2.2.1. Définition**

Soit  $X$  une variable aléatoire définie sur  $\Omega$ . On appelle fonction de répartition, l'application  $F$  qui, à tout réel, associe la probabilité que la variable aléatoire  $X$  prenne une valeur inférieure ou égale à  $x$ .

Pour tout réel  $x$ , on a :

$$F(x) = P(X \leq x)$$

$$P(a < X \leq b) = F(b) - F(a)$$

**NB :** cette notation relève de la littérature anglo-saxonne, la définition française donne par contre:

$$F(x) = P(X \geq x)$$

$$P(a \leq X < b) = F(b) - F(a)$$

**2.2.2. Propriétés**

- $F$  est une fonction croissante ;
- Pour tout  $x$ , on a  $0 \leq F(x) \leq 1$ , puisque  $F(x)$  est la probabilité de réaliser un événement ;
- $\lim_{x \rightarrow +\infty} F(x) = 1$
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- Pour tous réels  $a$  et  $b$   $P(a < X \leq b) = F(b) - F(a)$

**Remarque :** si l'on connaît la loi (ou la distribution) de la variable aléatoire  $X$ , on peut déterminer sa fonction de répartition  $F$  et réciproquement.

**Exemple 22 :** on jette successivement et indépendamment sur une table deux dés dont les faces sont numérotées de 1 à 4.  $X$  étant le produit des nombres obtenus.

- 1) Donner la loi de probabilité de  $X$ , son diagramme en bâtons.
- 2) Donner la fonction de répartition de  $X$  et sa courbe cumulative
- 3)

**2.3. Variables aléatoires indépendantes**

**2.3.1. Définition**

Deux variables aléatoires  $X$  et  $Y$  sont indépendantes si et seulement si, pour tous les couples  $(x, y)$  de valeurs possibles pour  $X$  et  $Y$ , les événements  $(X = x)$  et  $(Y = y)$  sont indépendants, ce qui s'écrit :

$$P(X = x \text{ et } Y = y) = P(X = x) \times P(Y = y)$$

De façon équivalente,  $X$  et  $Y$  sont indépendantes si et seulement si pour tous les réels  $x$  et  $y$  on a :

$$P(X \leq x \text{ et } Y \leq y) = P(X \leq x) \times P(Y \leq y)$$

**Exemple 23 :**

$x$	0	1	2
$P(X = x)$	0,6	0,3	0,1

$y$	1	2	3
$P(Y = y)$	0,6	0,3	0,1

$X$  et  $Y$  sont indépendantes. Dresser le tableau donnant les probabilités  $(X = x \text{ et } Y = y)$



Il s'agit d'un tableau à double entrées obtenu en utilisant  $P(X = x \text{ et } Y = y) = P(X = x) \times P(Y = y)$

#### 2.4. Caractéristiques d'une variable aléatoire discrète

Soit  $X$  une variable aléatoire définie sur  $\Omega$ .

##### 2.4.1. Espérance mathématique

L'espérance mathématique ou la moyenne d'une v.a est une mesure de tendance centrale. Elle est une moyenne pondérée des valeurs que peut prendre la v.a. Les poids correspondent aux probabilités. L'expression mathématique de l'espérance d'une variable aléatoire discrète  $X$  notée  $E(X)$  ou  $\bar{X}$  est définie par :

$$E(X) = \sum_{x \in X(\Omega)} X P(X = x)$$

**Remarque :** en particulier on a :

$$E(X^2) = \sum_{x \in X(\Omega)} X^2 P(X = x)$$

##### 2.4.1.1. Propriétés de l'opérateur espérance mathématique

$X$  étant une variable aléatoire,  $a$  et  $b$  deux constantes la propriété fondamentale de l'opérateur espérance mathématique est la linéarité. Ainsi on a :

- ✓  $E(a) = a$
- ✓  $E(aX) = aE(X)$
- ✓  $E(aX + b) = aE(X) + b$
- ✓  $E(X + Y) = E(X) + E(Y)$
- ✓ Inégalité de Jensen : soit  $g$ , une fonction non linéaire de  $x$  :
  - Si  $g$  est une **fonction convexe**:  $E[g(X)] \geq g[E(X)]$
  - Si  $g$  est une **fonction concave**:  $E[g(X)] \leq g[E(X)]$

##### 2.4.2. Variance et écart type

###### 2.4.2.1. Variance

Alors que l'espérance mathématique fournit la valeur moyenne de la variable aléatoire, on a souvent besoin d'une mesure de dispersion ou de variabilité. Si en statistique descriptive la variance est souvent utilisée pour comparer la dispersion de plusieurs variables, elle est ici utilisée pour résumer la dispersion des valeurs d'une v.a.

- **Définition**

La variance est une somme pondérée des écarts au carré d'une v.a par rapport à sa moyenne. Les pondérations correspondent aux probabilités.

L'expression mathématique de la variance d'une v.a discrète  $X$  notée  $V(X)$  est :

$$V(X) = \sigma^2 = \sum [X - E(X)]^2 P(X = x)$$

- **Théorème : (Koenig)**

Soit  $X$  une v.a discrète possédant une variance. Alors on a :

$$V(X) = E((X - \bar{X}))^2 = E(X^2) - E^2(X) = \sum_{x \in X(\Omega)} x^2 P(X = x) - E^2(X)$$

- **Propriété de la variance**

Soient la variable aléatoire  $X$ ,  $a$  et  $b$  deux constantes, on a :

- ✓  $V(X + a) = V(X)$

- ✓  $V(aX) = a^2V(X)$
- ✓  $\sigma(aX + b) = |a|\sigma(X)$
- ✓ La variance d'une somme de deux variables aléatoires est :  $V(X + Y) = V(X) + V(Y) - 2cov(X, Y)$
- ✓ Si X et Y sont indépendants,  $cov(X, Y) = 0$ ,  $V(X + Y) = V(X) + V(Y)$

#### 2.4.2.2. Ecart type

L'écart type de X noté  $\sigma(X)$  est la racine carrée de la variance

$$\sigma(X) = \sqrt{V(X)}$$

**Remarque :** l'écart type présente l'avantage d'être plus facile à interpréter que la variance puisqu'il est mesuré dans les mêmes unités que les données.

**Exemple 24 :** considérant l'exemple 21, calculez :

- a)  $E(X)$
- b)  $V(X)$
- c)  $\sigma(X)$

#### 2.4.3. Covariance de deux variables définies sur un même univers

La covariance de X et Y, notée  $cov(X, Y)$  est une mesure de la relation linéaire entre les deux variables aléatoires. Elle sert donc à évaluer la robustesse de la relation linéaire entre X et Y. son expression mathématique revient à calculer l'espérance de la v.a  $(X - \bar{X})(Y - \bar{Y})$ .

$$cov(X, Y) = E[(X - \bar{X})(Y - \bar{Y})]$$

$$cov(X, Y) = E(XY) - E(X)E(Y) = \sum \sum xyP(X = x \text{ et } Y = y) - E(X)E(Y)$$

#### 2.4.4. Propriétés de la covariance

- ✓  $cov(X, k) = 0$  et  $cov(k, Y) = 0$  k étant une constante donnée.
- ✓  $cov(X, Y) = cov(Y, X)$
- ✓  $Cov(aX + b, eY + d) = aeCov(X, Y)$ ; a, b, c et e sont des constantes

**Exemple 25 :** Déterminer l'espérance mathématique et la variance de l'exemple 18

#### 2.4.5. Moments non centrés et centrés

On appelle moment non centré d'ordre  $k \in \mathbb{N}^*$  l'espérance de la variable aléatoire  $X^k$  la quantité lorsqu'elle existe :

$$m_k = \sum X^k P(X = x) = E(X^k)$$

Le moment centré d'ordre  $k \in \mathbb{N}^*$  est :

$$\mu_k = \sum (X - \bar{X})^k P(X) = E[X - E(X)]^k$$

### III. VARIABLES ALEATOIRES CONTINUES

On dit qu'une variable aléatoire notée X est continue lorsque son ensemble de définition est un intervalle.

#### 3.1. Fonction de répartition

La fonction de répartition est le moyen le plus simple de caractériser la loi de probabilité d'une variable aléatoire.

### 3.1.1. Définition

On appelle fonction de répartition de la variable aléatoire  $X$  la fonction :

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

$$x \mapsto F_X(x) = P(X \leq x)$$

L'application  $F$  qui, à tout réel, associe la probabilité que la variable  $X$  prenne une valeur inférieure ou égale à  $x$  est par définition la fonction de répartition.

### 3.1.2. Propriétés

➤  $F$  est une fonction croissante, continue à droite, telle que

$$\lim_{x \rightarrow +\infty} F(x) = 1 \quad \lim_{x \rightarrow -\infty} F(x) = 0$$

➤ Pour tout  $x$ , on a  $0 \leq F(x) \leq 1$ , puisque  $F(x)$  est la probabilité de réaliser un événement. Elle permet de calculer la probabilité que  $X$  appartienne à n'importe quel intervalle de  $\mathbb{R}$  ;

➤ Pour tous réels  $a$  et  $b$ ,  $a < b$ , on a :  $P(a < X \leq b) = F(b) - F(a)$

**Exemple 26 :** Dans une banque, entre 9H et 10H le matin, le temps d'attente (en minute) de l'employé entre deux clients est une variable aléatoire dont la fonction de répartition est la fonction  $F$  définie par :

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

Un client vient juste de partir. Déterminer la probabilité que le client suivant se présente :

- Dans 5 minutes exactement ;
- Dans moins de 3 minutes ;
- Dans plus de 3min mais moins de 10 minutes

## 3.2. Loi de densité

### 3.2.1. Définition :

Si  $X$  est une variable aléatoire continue de fonction de répartition  $F$  dérivable, la loi de probabilité de  $X$  est définie par la dérivée de  $F$ , que nous notons  $f$ .

$f$  est appelé densité de probabilité ou fonction de densité de  $X$

$$f(x) = F'(x)$$

**Exemple 27 :** déterminer la fonction de densité de l'exemple 21

### 3.2.2. Propriétés

La densité de probabilité est une fonction croissante, sa dérivée est donc positive ou nulle

$$P(a < X \leq b) = \int_a^b f(t) dt = F(b) - F(a)$$

$$\int_{-\infty}^{+\infty} f(t) dt = 1$$

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

**3.2.3. Cas où X prend ses valeurs dans l'intervalle I de bornes  $\alpha$  et  $\beta$**

La densité est alors définie sur I seulement et l'on a :  $\int_{\beta}^{\alpha} f(t)dt = 1$ .

La fonction de répartition, qui est toujours définie sur R est donnée par :

- $F(x) = 0$  si  $x < \alpha$
- $F(x) = \int_{\alpha}^x f(t)dt$  si  $\alpha \leq x \leq \beta$
- $F(x) = 1$  si  $x > \beta$

**Exemple 28:** la quantité de café (en kg) vendu par semaine par la société Café Gbako est une variable aléatoire prenant ses valeurs dans l'intervalle [10 ; 70] de densité f définie par :

$$\begin{cases} f(x) = \frac{(10 - x)(x - 70)}{36000} & \text{si } 10 \leq x \leq 70 \\ f(x) = 0 & \text{sinon} \end{cases}$$

- a) Démontrer que f est une densité de probabilité ;
- b) Déterminer et représenter la fonction de répartition de X
- c) Calculer la probabilité que ce dernier vende la semaine prochaine :
  - Moins de 20 kg
  - Plus de 30 kg
  - Entre 20 et 30 kg

**3.3.Caractéristiques d'une variable aléatoire**

Soit X une variable aléatoire de densité f, prenant ses valeurs dans l'intervalle I de borne  $\alpha$  et  $\beta$  (ces bornes ou l'une d'elles étant éventuellement infinie)

**3.3.1. Espérance mathématique**

L'espérance mathématique est définie par :  $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$

**3.3.2. Variance**

La variance de X est :  $V(X) = \int_{\alpha}^{\beta} (t - E(X))^2 f(t)dt$

**3.3.3. Ecart type**

Ecart type de X est :  $\sigma(X) = \sqrt{V(X)}$

**3.3.4. Propriétés**

On retrouve les mêmes propriétés que dans le cas discret.

**Exemple 28:** Calculer l'espérance mathématique, la variance et l'écart type de la variable X de l'exemple 22

**IV. CONCLUSION**

Ce chapitre nous permis d'aborder les notions de bases sur les variables aléatoires discrètes et continues. Toutefois les notions telles que la notion de fonction génératrice d'une variables aléatoires, les démonstrations mathématiques de bon nombres de caractéristiques da variables aléatoires étudiées dans ce cours n'ont pas été abordé. Le lecteur intéressé par ces sujets pourra se référer aux multiples bibliographies qui existent dans le domaine. Le chapitre qui suit introduits les lois usuelles de probabilité les plus courantes.



# C

## HAPITRE III

# LOIS USUELLES DE PROBABILITES

Certaines situations expérimentales donnent des distributions de probabilités spécifiques. En économie, les distributions utilisées sont des modèles des phénomènes observés. Nous distinguons deux types de lois usuelles de probabilités qui sont les lois fondamentales discrètes et les lois fondamentales continues de probabilité. La première partie présente les différentes lois discrètes de probabilité et la deuxième les lois continues de probabilité.

### I. LES LOIS FONDAMENTALES DISCRETES DE PROBABILITES

#### 1.1. Loi de Bernoulli

##### 1.1.1. Modèle de base : Epreuve de Bernoulli

Il s'agit des variable aléatoires qui correspondent à la formalisation d'une expérience aléatoire qui ne peut avoir que deux issues possibles : succès/ échec ; pile/face ; répondre avoir l'intention de voter pour le candidat A/ répondre avoir l'intention de voter pour un autre candidat etc. Un des événements, en général le « succès » (c'est-à-dire l'événement qui nous intéresse en premier), est codé 1 et l'autre 0.  $\Omega = \{E, S\}$

- si  $S$  est réalisé,  $X = 1$
- si  $E$  est réalisé,  $X = 0$

##### 1.1.2. Définition

On appelle **variable de Bernoulli** ou variable *indicatrice*, la variable aléatoire  $X$  telle que :

$$X: \Omega \rightarrow \mathbb{R}$$

$$x \mapsto \Omega(X) = \{0,1\}$$

**Exemple 29 :** Lors d'un examen, un étudiant a une proportion  $p$  (avec  $0 < p < 1$ ) de succès et une proportion  $q$  (avec  $p + q = 1$ ) d'échec. Soit  $X$  la variable aléatoire « succès ». Les valeurs possibles de  $X$  sont 0 et 1 :  $X(\Omega) = \{0,1\}$  et les probabilités sont :

$$P(X = 1) = p$$

$$P(X = 0) = q$$

Avec  $p + q = 1$

On dit que la variable aléatoire  $X$  suit une loi de Bernoulli de paramètre  $p$ .

On l'écrit  $X \sim \mathfrak{B}(1, p)$ .

### 1.1.3. La fonction de répartition

Elle est définie par :

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ q & \text{si } 0 < x \leq 1 \\ 1 & \text{si } 1 < x \end{cases}$$

### 1.1.4. Valeurs caractéristiques (paramètres)

- ✓ Espérance mathématique est  $E(X) = p$
- ✓ La variance  $V(X) = pq$
- ✓ Ecart type  $\sigma(X) = \sqrt{pq}$

## 1.2. Loi binomiale

### 1.2.1. Définition

Soit une épreuve de Bernoulli répétée  $n$  fois ( $n > 1$ ), de manière *identiques et indépendantes* (la probabilité de succès reste donc  $p$  à chaque fois).

$$S_n: \Omega^n \rightarrow \mathbb{R}^n \\ x \mapsto S_n = X_1 + X_1 + \dots + X_i + \dots + X_n$$

Où  $X_i$  est une variable de Bernoulli.

Ainsi la loi de probabilité suivie par **la somme de  $n$  variables de Bernoulli** où la variable aléatoire  $S_n$  égale au nombre de succès obtenus au cours de  $n$  épreuves suit alors, par définition, **loi binomiale** de paramètres  $n$  et  $p$ . On note  $S_n \sim \mathcal{B}(n, p)$ .

$$S_n = \sum_{i=1}^n X_i \sim \mathcal{B}(n, p)$$

La probabilité que la variable binomiale  $S_n$  prenne la valeur  $k$  est donnée par la fonction de densité :

$$f(x) = P(S_n = k) = C_p^n p^k (1-p)^{n-k}, k \in \{0, 1, 2, \dots, n\}$$

### 1.2.2. Valeurs caractéristiques (paramètres)

Si  $X$  suit la loi binomiale  $\mathcal{B}(n, p)$

- $E(X) = pq$
- $V(X) = npq$
- $\rho(X) = \sqrt{npq}$

### 1.2.3. Lecture de la table associée à une loi binomiale notée $\mathcal{B}(n, p)$ pour $n \leq 10$

Pour une variable aléatoire  $X \sim \mathcal{B}(n, p)$ , la table donne les  $C_p^n p^k (1-p)^{n-k}$

- Pour des valeurs de  $n$  entre 2 et 10
- Des valeurs de  $k$  variant entre 0 et  $n$
- Des valeurs de  $p$  variant de 0,05 en 0,5 et se situant entre 0,05 et 0,5

**Exemple 30** : Soit  $X \sim \mathcal{B}(5; 0,3)$ , Calculer  $P(X \leq 2)$ .

Nous savons que :  $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

Pour  $n = 5$  et  $p = 0,3$  on lit sur la table de loi de binôme les valeurs respectives de :

- ✓  $P(X = 0) = 0,1681$
- ✓  $P(X = 1) = 0,3602$  et
- ✓  $P(X = 2) = 0,3087$

Nous avons donc :  $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0,837$

#### 1.2.4. Comportement asymptotique de la variable binomiale

Lorsque  $n > 10$ , il existe trois possibilités pour les calculs de probabilités de la variable binomiale:

○ *Approximation de la loi binomiale par une loi de Poisson ;*

Quand  $n \rightarrow +\infty$ ,  $p \rightarrow 0$ , de telle sorte que  $np \rightarrow m$ , fini, alors la loi binomiale tend vers la loi de poisson. Pour  $n \geq 50$  et  $p \leq 0,1$ , la loi binomiale  $B(n; p)$  peut être approchée par la loi de Poisson  $\mathcal{P}(np)$ .

○ **Approximation de la loi binomiale par une loi normale ;**

Quand  $n \rightarrow +\infty$ ,  $p$ , reste constant, alors la loi binomiale tend vers la loi normale.

○ *Autres cas*

Calculer les  $C_p^n p^k (1-p)^{n-k}$  manuellement dans les autres cas

**Exemple 31** : Grâce aux tarifs réduit de toutes sortes, 90% des voyageurs de la compagnie de transport *Karinou Airlines* et frère bénéficient de tarifs réduits. Chaque soir, un contrôleur prend les billets de 5 passagers choisis au hasard. Déterminer sa loi de probabilité, calculer son espérance mathématique

#### 1.2.5. Somme de variables aléatoires binomiales indépendantes

Soient  $X_1$  et  $X_2$  deux variables aléatoires indépendantes suivant respectivement les lois binomiales  $\mathcal{B}(n_1, p)$  et  $\mathcal{B}(n_2, p)$ , alors  $X_1 + X_2$  suit la loi binomiale de paramètre

$n = n_1 + n_2$  et d'espérance mathématique  $p$ . On note :  $(X_1 + X_2) \sim \mathcal{B}(n_1 + n_2, p)$

### 1.3. Loi hypergéométrique

La loi hypergéométrique est étroitement liée à la loi binomiale. La différence majeure entre les deux lois est que, lorsqu'il s'agit d'une loi hypergéométrique, les tirages ne sont pas indépendants, et la probabilité de succès change d'un tirage à l'autre.

#### 1.3.1. Définition

Soit une urne contenant  $N$  boules dont :

- une population  $N_p$  est constituée de boules de couleur blanche ( $p$  proportion de boules blanches, événement succès).
- une population  $N_q$  est constituée de boules de couleur noire ( $q$  proportion de boules noire, événement échec).

On prélève dans cette urne un échantillon sans remise de  $n$  ( $n \leq N$ ) boules. La composition de l'urne varie ainsi à chaque tirage, de ce fait les tirages ne sont pas indépendants.

Soit  $X$  le nombre de boules blanches de l'échantillon.

Soit la variable aléatoire  $X$  égale au nombre de boules blanches tirées (nombre de succès obtenus) au cours de  $n$  épreuves. La variable aléatoire  $X$  suit alors, par définition, la loi hypergéométrique de paramètres  $N, n$  et  $N_p$ .

On note :  $X \sim \mathcal{H}(N, n, N_p)$ .

La probabilité que la variable hypergéométrique  $X$  prenne la valeur  $k$  (nombre de boules blanches tirées) est donnée par la fonction densité :

$$f(x) = P(X = k) = \frac{\binom{k}{N_p} \binom{n-k}{N_q}}{\binom{n}{N}}$$

avec  $\max(0, n - (N - N_p)) \leq k \leq \min(n, N_p)$

$N_q$ : Effectif des boules de couleur non blanche.

**Exemple 32** : un jeu de cartes contient 52 cartes, dont 4 as. Quelle est la probabilité que la donne de cinq cartes fournisse :

- une paire d'as
- un as
- aucun as
- au moins un as

### 1.3.2. Valeurs caractéristiques (paramètres)

Si  $X$  suit la loi hypergéométrique  $X \sim \mathcal{H}(N, n, N_p)$

- $E(X) = np$
- $V(X) = npq \frac{N-n}{N-1}$

**NB** : le coefficient  $\frac{N-n}{N-1} \leq 1$  qui diffère peu de  $1 - \frac{n}{N}$  lorsque  $N$  est grand, est appelée *coefficient d'exhaustivité*.

**Exemple 33**: Une urne contient 10 boules dont 6 blanches et 4 noires.

On tire en une seule fois trois boules de l'urne (tirage exhaustif d'effectif trois). Soit  $Y$  le nombre d boules blanches extraites.

- Donner numériquement la loi de probabilité de  $Y$ .
- Indiquer  $E(Y)$
- Indiquer  $V(Y)$

### 1.3.3. Comportement asymptotique de la variable hypergéométrique : Convergence de la loi hypergéométrique vers la loi binomiale

Si  $N$  tend vers l'infini, la loi  $\mathcal{H}(N, n, N_p)$  tend vers la loi  $B(n, p)$ , c'est-à-dire que lorsqu'on effectue un tirage dans une grande population, il importe peu que ce tirage se fasse avec ou sans remise (en pratique, on considèrera que la population est « grande » lorsque l'échantillon représente moins de

10% de cette population :  $\frac{n}{N} < 0,1$ ).

## 1.4. Loi de Poisson

La **loi de Poisson** découverte au début du XIX<sup>e</sup> siècle par le magistrat français **Siméon-Denis Poisson** s'applique souvent aux phénomènes accidentels où la probabilité  $p$  est très faible ( $p < 0,05$ ). Elle peut également dans certaines conditions être définie comme **limite d'une loi binomiale**.

### 1.4.1. Génération de la loi de Poisson: Processus poissonnien

On appelle **processus poissonnien** (ou processus de Poisson), le modèle probabiliste des situations qui voient un flux d'évènements se produire les uns à la suite des autres de façon aléatoire (dans le temps et dans l'espace), obéissant aux conditions suivantes :

- 1) la probabilité de réalisation de l'évènement au cours d'une petite période ou sur une petite portion d'espace  $\Delta t$  est proportionnelle à  $\Delta t$  soit  $p\Delta t$ .
- 2) elle est indépendante de ce qui s'est produit antérieurement ou à côté,
- 3) la probabilité de deux apparitions sur le même  $\Delta t$  est négligeable.

Ainsi, des évènements qui se réalisent de façon aléatoire comme :

- nombre de fautes dans un texte,
- Nombre de pièces défectueuses dans un lot ;
- Nombre d'erreur dans une comptabilité ;
- Nombre de pannes d'un appareil au cours d'une période donnée ;
- Nombre d'appels téléphoniques reçu au cours d'une période donnée ;
- Nombre de personnes arrivant à un guichet au cours d'une période donnée (arrivées dites poissonnières)

peuvent être considérés comme relevant d'un processus poissonnier. C'est pourquoi cette loi est également appelée loi des petites probabilités ou loi des « phénomènes rares », elle se rencontre souvent pour des phénomènes accidentels.

#### 1.4.2. Définition :

Soient  $\lambda$  un réel strictement positif donné et  $X$  une variable aléatoire.

$X$  suit la loi de poisson de paramètre  $\lambda$  si elle peut prendre toutes les valeurs entières  $k$  avec la probabilité :

$$f(x) = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

On note  $X \sim \mathcal{P}(\lambda)$

#### 1.4.3. Valeurs caractéristiques (paramètres)

- $E(X) = \lambda$
- $V(X) = \lambda$
- $\rho(X) = \sqrt{\lambda}$

**Exemple 34 :** considérer une distribution de probabilité de poisson avec  $\lambda = 3$

- a) Ecrire la fonction de probabilité de poisson appropriée.
- b) Calculer  $f(2)$
- c) Calculer  $f(1)$
- d) Calculer  $P(x \geq 2)$

#### 1.4.4. Somme de variables aléatoires de poisson indépendantes

Si  $X_1$  et  $X_2$  deux variables aléatoires indépendantes suivant respectivement les lois de Poisson  $\mathcal{P}(\lambda_1)$  et  $\mathcal{P}(\lambda_2)$ , alors  $X_1 + X_2$  suit la loi de Poisson de paramètre  $\lambda = \lambda_1 + \lambda_2$ . On note :  $(X_1 + X_2) \sim \mathcal{P}(\lambda_1 + \lambda_2)$

#### 1.4.5. Lecture de la table associée à une loi de Poisson

La table associée à une loi de Poisson donne un certain nombre de valeur du paramètre  $\lambda$  en lignes et de valeur de  $k$  données en colonnes.

**Exemple 35:** Soit  $X \sim \mathcal{P}(3)$ , alors  $P(X = 2) = 0,2240$

#### 1.4.6. Approximation d'une loi binomiale par une loi de poisson

Si les trois conditions suivantes sont vérifiées :

- $n$  "grand" par exemple  $n \geq 30$
- $p$  "petite" (par exemple  $p \leq 30$ )

- Et  $np$  "pas trop grand" (par exemple  $np \leq 15$ )

On peut donc approcher la loi binomiale  $\mathcal{B}(n, p)$  par la loi de Poisson de paramètre  $\lambda = np$ , ce qui permet d'écrire pour  $k$  entier inférieur ou égal à  $n$  :

$$P(X = k) = C_p^n p^k (1 - p)^{n-k} \sim e^{-np} \frac{(np)^k}{k!}$$

## II. LOIS FONDAMENTALES CONTINUES DE PROBABILITES

### 2.1. Loi uniforme

#### 2.1.1. Définition

Une variable aléatoire  $X$  suit la loi uniforme continue si sa fonction de densité est constante sur un intervalle fini  $[a, b]$ , s'écrivant :

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

On note :  $X \sim \mathcal{U}([a, b])$

On l'appelle également loi de rectangle compte tenu du graphe de sa fonction de densité.

On sait que  $\int_{-\infty}^{+\infty} f(t) dt = 1 = \int_a^b k dt$ .

#### 2.1.2. fonction de répartition

La fonction de répartition, qui est toujours définie sur  $\mathbb{R}$  est donnée par :

$$F(X) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases}$$

#### 2.1.3. Espérance mathématique, variance et écart type

- $E(X) = \frac{b+a}{2}$
- $V(X) = \frac{(b-a)^2}{12}$

### 2.2. Loi exponentielle

#### 2.2.1. Définition

Soit  $\alpha$  un réel strictement positif donné

Une variable aléatoire continue  $X$  suit la loi exponentielle de paramètre  $\alpha$  si elle admet pour de densité la fonction  $f$  de  $\mathbb{R}$  dans  $\mathbb{R}$  définie par :

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

On écrit  $X \sim \mathcal{E}(\theta)$

La loi exponentielle est souvent utilisée pour représenter une durée de vie (durée de chômage, durée de vie d'un matériel, durée d'hospitalisation)

#### 2.2.2. Fonction de répartition

On sait que la fonction de répartition  $F$  d'une variable aléatoire continue de densité  $f$  vérifie :

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \theta \int_0^x e^{-\theta t} dt = [-e^{-\theta t}]_0^x$$

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\mu x} & \text{si } x \geq 0 \end{cases}$$

### 2.2.3. Espérance mathématique, variance et écart type

Soit  $X$  une variable aléatoire suivant la loi exponentielle de paramètre  $\alpha$ .

- Son espérance mathématique est :  $E(X) = \frac{1}{\theta}$
- Sa variance est :  $V(X) = \frac{1}{\theta^2}$
- Son écart type est :  $\rho(X) = \frac{1}{\theta}$

**Exemple 36:** on considère un composant électronique dont la durée de vie suit une loi exponentielle d'espérance mathématique 10 ans. Déterminer la densité, la fonction de répartition et l'écart type. Calculer la probabilité qu'un tel composant ait une durée de vie :

- Supérieure à 5 ans
- Supérieure à 10 ans
- Comprise entre 5 et 10 ans

Un composant électronique de ce type fonctionne déjà depuis 8 ans. Quelle est la probabilité qu'il fonctionne encore au moins : 5 ans ? 10 ans ? Que remarque-t-on ?

## 2.3. Loi normale ou loi de Laplace-Gauss

Dans la pratique, la loi normale ou la loi de Laplace Gauss est une des distributions que l'on rencontre souvent. Nombreux phénomènes peuvent être représentés de façon satisfaisante par cette loi.

De façon générale, la loi normale modélisera les situations aléatoires possédant de nombreuses causes indépendantes dont les effets s'ajoutent, sans que l'un d'eux soit prépondérant (qualité d'une production industrielle, erreurs de mesure, gestion de comptes bancaires...).

### 2.3.1. Définition de la loi normale

Soit  $m$  un réel quelconque et  $\sigma > 0$ . Une variable aléatoire continue  $X$  suit la loi normale d'espérance  $m$  et d'écart type  $\sigma$  si elle admet pour densité la fonction  $f$  définie sur  $\mathbb{R}$  par

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

$m$  et  $\sigma > 0$  sont respectivement les paramètres de la loi. On note  $X \sim \mathcal{N}(m, \sigma)$ .

### 2.3.2. fonction de répartition

la fonction de répartition représentant la probabilité que la variable aléatoire  $X$  ait une probabilité inférieure ou égale à  $x$ , a pour expression :

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2} \frac{(t-m)^2}{\sigma^2}} dt$$

### 2.3.3. Espérance mathématique, variance et écart type

Soit  $X$  une variable aléatoire suivant la loi normale de paramètre  $m, \sigma$ .

- Son espérance mathématique est :  $E(X) = m$
- Sa variance est :  $V(X) = \sigma^2$
- Son écart type est :  $\sigma(X) = \sigma$

### 2.4. Loi normale centrée réduite

Pour effectuer des calculs de probabilités, l'expression mathématique de la loi normale est d'utilisation malaisée. Soit  $T$  une variable aléatoire. On effectue un changement de variable en posant  $T = \frac{X-m}{\sigma}$ , il s'en suit que  $T$  est normée si :

- ✓ son espérance mathématique est nulle :  $E(X)=0$  (on dit que la v.a est centrée)
- ✓ son écart type est un,  $\sigma(X) = 1$ . (on dit que la v.a est réduite)

Ainsi  $T$  obéit à une loi normale centrée et réduite.

#### 2.4.1. Définition

Une variable aléatoire continue  $T$  suit la loi normale centrée réduite si elle admet pour densité la fonction  $f$  définie sur  $\mathbb{R}$  par :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

On note  $T \sim \mathcal{N}(0, 1)$

Le changement de variable permet de ramener n'importe quelle distribution normale à une même loi de probabilité normale centrée réduite pour laquelle nous disposons de tables numériques.

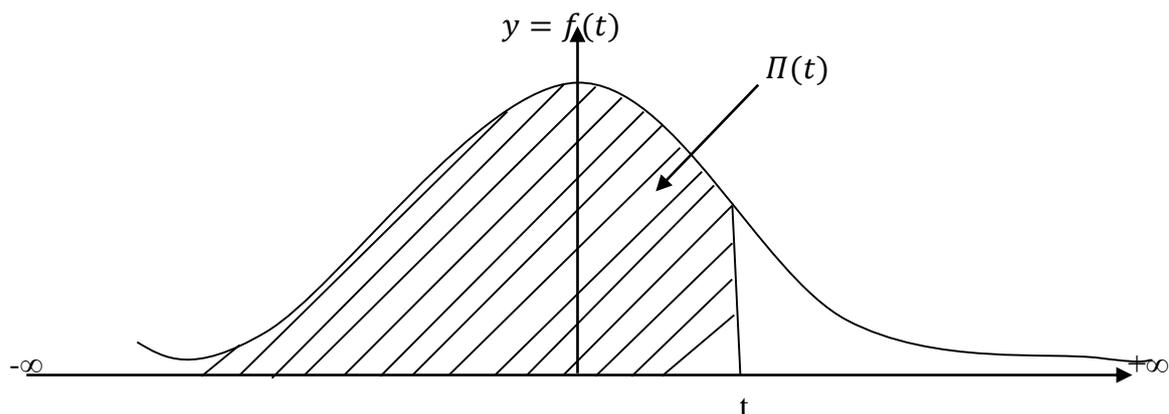
#### 2.4.2. Fonction de répartition

Si  $T$  est une variable aléatoire continue suivant la loi normale centrée réduite, sa fonction de répartition, souvent notée  $\Pi$  est définie par :

$$\begin{aligned}\Pi(t) &= P(T \leq t) = \int_{-\infty}^t f(x) dx \\ \Pi(t) &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx\end{aligned}$$

- **Représentation graphique :**

$\Pi(t)$  est l'aire hachurée ci-dessous



#### 2.4.3. Propriétés :

1.  $\Pi(-\infty) = 0$  et  $\Pi(+\infty) = 1$
2.  $\Pi(0) = P(T \leq 0) = 0,5$
3.  $\Pi(-t) = 1 - \Pi(t)$  pour tout  $t$
4. La fonction  $\Pi$  est strictement croissante
5. Nous avons aussi  $P(a \leq T \leq b) = \Pi(b) - \Pi(a)$

#### 2.4.4. Espérance, variance et écart type

Soit  $X$  une variable aléatoire suivant la loi normale centrée réduite de paramètre  $m, \sigma$ .

- ✓ Son espérance mathématique est :  $E(X) = 0$
- ✓ Sa variance est :  $V(X) = 1$
- ✓ Son écart type est :  $\sigma(X) = 1$

#### 2.4.5. Lecture de tables de la loi normale centrée et réduite

Deux tables relatives à la loi normale centrée réduites sont utilisées. La table de la fonction de répartition (ou fonction intégrale) et la table de l'écart réduit.

#### 2.4.6. Table de la fonction de répartition de la loi normale centrée réduite

Pour  $t > 0$ , on peut lire la valeur de  $\Pi(t)$  dans la table

Utiliser la table de la fonction intégrale de la loi normale centrée réduite pour déterminer :  $\Pi(1)\Pi(2,3)$ ,  $\Pi(-1,92)$ . Le réel  $a$  tel que  $\Pi(a) = 0,975$  le réel  $b$  tel que  $\Pi(b) = 0,006$

**Exemple 37:** Soit  $T$  une variable aléatoire continue de loi normale centrée réduite. Déterminer les probabilités suivantes :

- $P(-1 < T < 2)$ ,  $P(-1 \leq T < 1)$
- Déterminer  $h$  tel que  $P(|T| \leq h) = 0,95$
- Exprimer en fonction de  $\Pi(t)$  la probabilité  $P(|T| > t)$

#### 2.4.7. Table de l'écart réduit

Pour  $\alpha \in ]0,1[$ , on peut trouver  $t > 0$  tel que :  $P\{-t \leq N(0,1) \leq t\} = 1 - \alpha$

- Si  $\alpha = 0,01$  on lit sur la table de l'écart réduit  $t =$
- Si  $\alpha = 0,05$  on lit sur la table de l'écart réduit  $t =$
- Si  $\alpha = 0,10$  on lit sur la table de l'écart réduit  $t =$

#### 2.4.8. Opérations sur les variables aléatoires normales

##### 2.4.8.1. Fonction affine d'une variable normale

$a$  et  $b$  étant deux réels donnés, et  $X$  une variable aléatoire qui suit la loi normale de paramètre  $m$  et  $\sigma$ , la variable  $Y = aX + b$  suit la loi normale d'espérance  $am + b$  et d'écart type  $|a|\sigma$

##### 2.4.8.2. Somme et différence de deux variables aléatoires normales indépendantes

Si  $X_1$  est une variable qui suit la loi normale de paramètre  $m_1$  et  $\sigma_1$  ;

Si  $X_2$  est une variable qui suit la loi normale de paramètre  $m_2$  et  $\sigma_2$  ;

Et si  $X_1$  et  $X_2$  sont indépendants, alors

$X_1 + X_2$  suit la loi normale d'espérance  $m_1 + m_2$  et d'écart type  $\sqrt{\sigma_1^2 + \sigma_2^2}$

**Exemple 38:** la société d'exploitation forestière THANRY SA dispose de deux ateliers. On admettra que la production de ces deux ateliers est normale et indépendante. Le premier a en moyenne une production journalière de 2000 tonnes avec un écart type de 150 tonnes. Le deuxième atelier dispose d'une capacité de production journalière moyenne de 3000 tonnes avec un écart type de 200 tonnes. Quelle est la probabilité que la production journalière n'excède pas 5200 tonnes.

#### 2.4.9. Le théorème central limite

La loi normale est d'application générale. Elle est engendrée sous des conditions très peu restrictives par l'addition de causes de fluctuation nombreuses et indépendantes. Il s'agit donc du théorème central limite qui est un théorème très important du calcul des probabilités.

Le théorème s'énonce comme suit : Soient  $(X_1, X_2, \dots, X_n)$ , une suite de variables aléatoires indépendantes de même loi, d'espérance mathématique finie  $m$  et de variance finie  $\sigma^2$  et si  $n$  augmente indéfiniment quelques soit les lois de probabilités suivies par  $(X_n \text{ pour tout } n \in \mathbb{N}^*)$ , la somme des variables aléatoires  $S_n = \sum_{i=1}^n X_i$  suit sensiblement la loi normale de d'espérance  $nm$  et de variance  $n\sigma^2$ . Alors la variable aléatoire:  $Z_n = \frac{S_n - nm}{\sigma\sqrt{n}}$  converge vers une variable centrée réduite

**2.4.10. Approximations par une loi normale**

Au chapitre 2, nous avons vu que sous certaines conditions, la loi binomiale peut être approchée par une loi de Poisson. Cette section est consacrée à l'étude des situations permettant d'approcher une binomiale d'une part et d'autre par une loi de Poisson par une loi normale

**2.4.10.1. Approximation d'une loi binomiale par une loi normale**

Si la variable aléatoire discrète  $X$  suit la loi binomiale  $\mathcal{B}(n, p)$  et si l'une des conditions suivantes est réalisée :

- [n " grand " ( $n \geq 30$ ) ;  $p$  et  $q$  voisins de 0,5]
- Ou [ $np > 15$  et  $nq > 15$ ]
- Ou [ $npq > 10$ ]
- On peut donc approcher la loi  $X$  par la loi normale de paramètres :

D'espérance mathématique  $m = np$  et d'écart type  $\sigma = \sqrt{npq}$

On note  $X^*$  la variable aléatoire qui la loi  $\mathcal{N}(np, \sqrt{npq})$

**2.4.10.2. Approximation d'une loi de Poisson par une loi normale**

Si la variable aléatoire discrète  $X$  suit la loi de Poisson  $\mathcal{P}(\lambda)$  et si

$\lambda$  est « grand »  $\lambda > 10$  On peut approcher la loi  $X$  par la loi normale de paramètres :

D'espérance mathématique  $m = \lambda$  et d'écart type  $\sigma = \sqrt{\lambda}$

On note  $X^*$  la variable aléatoire qui la loi  $\mathcal{N}(\lambda, \sqrt{\lambda})$

**III. LOIS DE PROBABILITES DERIVEES**

**3.1. Loi Gamma**

**3.1.1. Définition**

Soit  $X$  une variable aléatoire qui suit la loi de gamma de paramètres  $p > 0$ , et  $\theta > 0$ ). La densité de probabilité est donnée par :

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

Cette loi est résumée par la notation  $\Gamma(p)$  avec :

$$\Gamma(p) = \int_0^{+\infty} t^{p-1} e^{-t} dt$$

On écrit  $X \sim \Gamma(p, \theta)$

La loi gamma est utilisée dans le cas de la distribution de revenu, fonction de production etc.

**3.1.2. Propriété de  $\Gamma(p)$**

- $\Gamma(p + 1) = p\Gamma(p)$
- $\Gamma(p) = (p - 1)!$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

### 3.1.3. Valeurs caractéristiques

- L'espérance mathématique de la loi gamma est :  $E(X) = \frac{p}{\theta}$
- La variance vaut :  $V(X) = \frac{p}{\theta^2}$

### 3.1.4. Somme de deux variables aléatoires indépendantes

Soient  $X$  et  $Y$  deux variables aléatoires indépendantes suivant respectivement les lois  $\Gamma(p, \theta)$  et  $\Gamma(q; \theta)$  :

$$X \sim \Gamma(p, \theta)$$

$$Y \sim \Gamma(q, \theta)$$

Alors la somme des variables aléatoires notée  $Z$  suit une loi gamma de paramètre  $q + p$

On note  $Z = X + Y \sim \Gamma(p + q, \theta)$

La famille des lois gamma contient comme cas particuliers deux lois usuelles :

- La loi exponentielle de paramètre  $\theta > 0$  obtenue pour  $p = 1$  et qui est donc la loi  $\Gamma(1, \theta)$  de densité  $f(x) = \theta e^{-\theta x}$  pour  $x > 0$
- La loi de khi-deux, très utilisée en statistique que nous verrons au point 3

## 3.2. Loi du khi – deux $\chi^2(n)$

### 3.2.1. Définition

Soit  $(X_1, X_2, \dots, X_n)$   $n$  variables indépendantes qui suivent toutes la loi normale centrée et réduite. La somme de leur carrée ( $X = X_1^2 + X_2^2 + \dots + X_n^2$ ) est une variable aléatoire qui suit la loi de khi-deux à  $n$  degré de liberté notée  $\chi^2(n)$  ou  $\chi_n^2$ . Sa fonction de densité est donnée par :

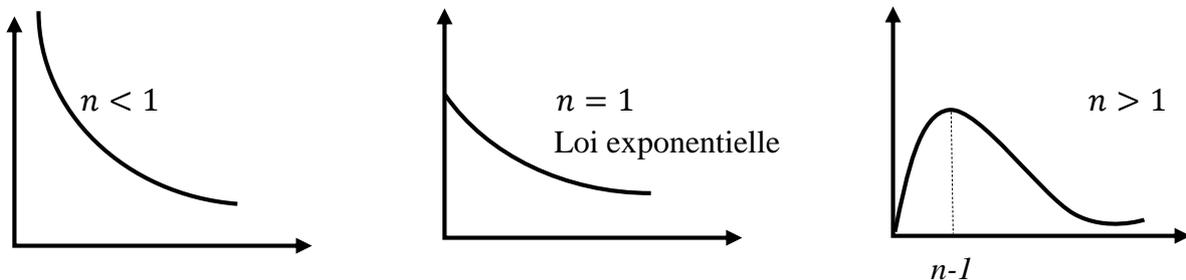
$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{1}{2}x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

### 3.2.2. Forme de la distribution

Pour  $n = 1$ , on obtient la loi  $\chi^2(1)$  appelé encore loi exponentielle.

La distribution du  $\chi^2$  est dissymétrique avec étalement vers la droite. Toutefois, lorsque le nombre de degré de liberté  $n$  augmente, la distribution tend à devenir symétrique : elle s'approche alors de la distribution normale à laquelle elle peut être approchée lorsque  $n > 30$

**Figure 4 :** Histogrammes de la loi  $\chi^2(n)$



**3.2.3. Valeurs caractéristiques**– **Espérance mathématique**

$$E(\chi^2(n)) = \frac{n/2}{1/2} = n$$

– **Variance**

$$V(\chi^2(n)) = \frac{n/2}{1/4} = 2n$$

**3.2.4. Somme de deux variables indépendantes de loi de Khi- Deux**

Soit  $X$  et  $Y$  deux variables aléatoires indépendantes suivant la loi de  $\chi^2$  qui ont respectif  $n_1$  et  $n_2$  degrés de liberté. Leur somme suit donc une loi de  $\chi^2$  à  $n_1 + n_2$  degré de liberté.

Si  $X \sim \chi^2(n_1)$  et  $Y \sim \chi^2(n_2)$ , deux variables indépendantes alors :

$$S = X + Y \sim \chi^2(n_1 + n_2)$$

Ce résultat se généralise au cas de  $m$  variables aléatoires indépendantes

**3.2.5. Table de la loi de khi-deux**

La loi de khi-deux ne dépend que d'un seul paramètre  $n$  le nombre de degré de liberté. La table d'une loi de khi-deux est une table à double entrée ( $n$  et  $\alpha$ ) qui donne pour  $n \leq 30$ , la valeur du  $\chi^2$  ayant la probabilité  $\alpha$  d'être dépassée.

$$P(\chi_n^2 > \chi_\alpha^2) = \alpha$$

Si les valeurs de  $n$  et  $\alpha$  sont données, on lit  $\chi_\alpha^2$  sur la table de khi-deux

**Exemple 39:** Loi de Khi-deux à 10 degrés de liberté

Pour  $\alpha = 0,05$  on lit sur la table  $\chi_\alpha^2 = 18,30$ .

Pour  $\alpha = 0,90$  on lit sur la table  $\chi_\alpha^2 = 4,86$ .

Pour  $n = 10$ , La valeur de  $\chi_\alpha^2$  a une probabilité de 5% d'être supérieure à 18,30 et une probabilité de 90% d'être supérieure à 4,86

**3.2.6. Approximation de la loi de khi-deux par la loi normale**

Soit  $X$  une variable aléatoire indépendante suivant une loi du khi-deux à  $n$  degrés de liberté.  $n > 30$ , la variable suit alors approximativement une loi normale centrée et réduite.

**3.3. Loi de Student****3.3.1. Définition**

Soient  $X$  et  $Y$  deux variables aléatoires indépendantes suivant respectivement la loi normale centrée et réduite, et la loi de khi-deux.  $X \sim \mathcal{N}(0, 1)$ ;  $Y \sim \chi_n^2$

On appelle loi de Student à  $n$  degrés de liberté notée  $T_n$ ; la loi suivie par le rapport :

$$T_n = \frac{X}{\sqrt{\frac{Y}{n}}}$$

Elle a pour densité :

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} ; x \in \mathbb{R}$$

### 3.3.2. Valeurs caractéristiques

#### - Espérance mathématique

$E(T_n) = 0$  du fait de la symétrie de la distribution

#### - Variance

$$V(T_n) = \begin{cases} \frac{n}{n-2} ; \text{si } n > 2 \\ \infty ; \text{si } n \leq 2 \end{cases}$$

La loi de Student est utilisée lorsqu'on veut vérifier l'hypothèse de la moyenne d'une population normalement répartie de variance inconnue. En outre, cette loi est utilisée souvent en économétrie pour tester l'hypothèse de la non significativité des variables indépendantes du modèle.

### 3.3.3. Comportement asymptotique

Cette loi a la même allure que la loi normale mais avec des queues plus affines. Elle tend asymptotiquement vers la loi normale centrée et réduite lorsque  $n > 30$ .

$$T_n \rightsquigarrow \mathcal{N}(0, 1), \text{ quand } n > 30$$

**Remarque :** Loi de Student à 1 degré de liberté

La loi de Student est confondue à celle de Cauchy lorsque  $n = 1$ . La densité de la loi de Cauchy est définie par :  $f(x) = \frac{1}{\pi(1+x^2)}$  ;  $x \in \mathbb{R}$

### 4.4.5. Table de loi de Student

La table de loi de Student donne la probabilité  $\alpha$  pour que  $T_n$  égale ou dépasse, en valeur absolue, une valeur donnée  $t_\alpha$ .

$$P(-t_\alpha \leq T_n \leq t_\alpha) = 1 - \alpha$$

Si les valeurs  $n$  et  $\alpha$  sont données, on lit alors  $t_\alpha$  sur la table de Student.

**Exemple 40:** la loi de Student à 15 degrés de liberté pour  $\alpha = 5\%$  et  $\alpha = 90\%$

Pour  $n = 15$  au seuil  $\alpha = 5\%$ , on lit  $t_\alpha = 2,13$  et  $t_\alpha = 0,128$  pour  $\alpha = 90\%$

## 4.5. Loi de Fischer – Snedecor

### 4.5.1. Définition

Soient  $X$  et  $Y$  deux variables indépendantes suivant les lois de khi-deux respectivement à  $n$  degrés de liberté  $[\chi^2(n)]$  et  $m$  degrés de liberté  $[\chi^2(m)]$ . Alors les rapports  $\frac{X/n}{Y/m}$  suit une loi de Fischer – Snedecor à  $n$  et  $m$  degré de liberté, notée  $F(n, m)$

Pour  $F > 0$ , la fonction de densité de probabilité est de la forme :

$$f(F) = C_{(n,m)} F^{\frac{n}{2}-1} (m + nF)^{-\left(\frac{n+m}{2}\right)}$$

$$F \leq 0, \quad f(F) = 0$$

### 4.5.1. Espérance et variance

- L'espérance de la variable de Fisher-Snédecor est :  $E(F) = \frac{m}{m-2}$ , si  $m > 2$
- La variance de la variable de Fisher-Snédecor est :  $V(F) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$ , si  $m > 4$

### 4.5.2. Table de la loi de Fisher – Snedecor

# S

## ECONDE PARTIE :

# CONCEPTS ET PROCEDURES DE L'INFERENCE STATISTIQUE

Pour étudier une population statistique on a recours à deux méthodes : la méthode exhaustive et la méthode des sondages. La méthode exhaustive encore appelé recensement consiste à examiner chaque individu de la population selon les caractères étudiés.

La seconde méthode dite de sondage consiste à examiner, observer ou interroger une partie de la population, c'est ce qu'on appelle enquête partielle ou sondage. Les éléments de la population qui sont réellement observés constituent l'échantillon et l'opération qui consiste à choisir ces éléments est appelée échantillonnage.

Par rapport à l'enquête complète, l'enquête partielle offre une série d'avantages :

1. Le coût global de l'enquête partielle est en général plus réduit que le coût global d'une enquête complète (rapport coût).
2. L'enquête par sondage est plus rapide que l'enquête complète, surtout lorsque la caractéristique étudiée présente des modifications assez importantes au cours du temps (rapport temps).
3. Les erreurs d'observations sont plus réduites que dans l'enquête exhaustive (qualité observations).
4. Enfin dans certaines situations particulières, l'enquête partielle est la seule solution possible, c'est le cas lorsque l'observation présente un caractère destructif.

L'utilisation des résultats obtenus sur une partie de la population pour estimer ou tester des hypothèses sur les caractéristiques de la population entière constitue l'inférence statistique. Les premiers problèmes d'inférence statistique auxquels s'applique la théorie des distributions d'échantillonnage seront alors les problèmes d'estimations. Le but poursuivi est d'estimer, à partir d'un échantillon, la ou les valeurs numériques d'un ou de plusieurs paramètres de la population considérée et de déterminer la précision de cette ou de ces estimations. Le second est celui des tests statistiques. Un test statistique est une méthode permettant de prendre une décision à partir d'informations fournies par un échantillon. Ces deux méthodes de sondages feront successivement l'objet des chapitres 5 et 6 de cette seconde partie, après une brève aperçue de la théorie d'échantillonnage dans le chapitre 4.

# C

## HAPITRE IV

# THEORIE D'ECHANTILLONNAGE

La théorie de l'échantillonnage porte sur l'étude des relations ou rapports existants entre une population (on dit parfois population-mère) et les échantillons prélevés dans cette population. L'objet de ce chapitre sera d'une part d'étudier les différentes méthodes de construction d'un échantillon et d'autre part on y aborde la théorie de la distribution d'échantillonnage à partir des exemples concrets.

### I. VOCABULAIRE

**Population** : ensemble de tous les éléments d'intérêt dans une étude particulière.

La population peut être finie (exemple nombre de ménages) ou infinie (exemple l'ensemble de nombre réels). Une population finie suffisamment grande, dans laquelle le tirage est effectué avec remise peut théoriquement être considérée comme infinie :

**Population FINIE + Tirage AVEC Remise = Population INFINIE.**

**Echantillon** : sous-ensemble de la population. De manière pratique c'est l'ensemble des unités de base sélectionnées et réellement observées au cours d'un sondage.

**Enquête** : ensemble des opérations de collecte et de traitement de données relatives à quelques domaines que ce soit.

**Recensement** : Enquête visant à collecter des données relatives à la population entière. (Enquête complète ou enquête exhaustive).

**Sondage** : Enquête visant à collecter des données relatives à un échantillon. (Enquête incomplète, enquête partielle ou enquête par échantillonnage).

**Unité de base** : unité d'échantillonnage ou unité de sondage, c'est l'élément pris en considération dans l'enquête.

**Echantillonnage** : ensemble des opérations qui permettent de sélectionner de façon organisée les éléments de l'échantillon.

**Cadre** : une liste d'éléments à partir desquels l'échantillon est sélectionné.

**Base de sondage** : énumération ou présentation ordonnée de toutes les unités de base constituant la population.

**Erreur d'échantillonnage** : écart entre les résultats obtenus auprès d'un échantillon et ce que nous apprendrait un recensement comparable de la population. Plus la taille de l'échantillon est grande plus l'erreur d'échantillonnage diminue.

**Fraction ou taux de sondage** : proportion des unités de la population qui font partie de l'échantillon. C'est le rapport entre la taille de l'échantillon  $n$ , et la taille de la population  $N$ .

$$f = n/N \times 100$$

## II. METHODES D'ECHANTILLONNAGE

Plusieurs méthodes peuvent être utilisées pour sélectionner un échantillon dans une population. On distingue les méthodes probabilistes où les éléments tirés de la population ont une probabilité connue de faire partie de l'échantillon et les méthodes dites non probabilistes. Au nombre des méthodes probabilistes on compte l'échantillonnage aléatoire simple, l'échantillonnage aléatoire stratifié, l'échantillonnage par grappes et l'échantillonnage systématique. Les méthodes non probabilistes sont l'échantillonnage de commodité et l'échantillonnage subjectif.

### 2.1.Méthodes d'échantillonnage probabiliste

#### 2.1.1. Echantillonnage Aléatoire Simple (EAS)

Un échantillonnage est aléatoire si tous les individus de la population ont la même probabilité de faire partie de l'échantillon, il est simple si les prélèvements des individus sont réalisés indépendamment les uns des autres.

Dans la pratique, la procédure de sélection d'un tel échantillon varient selon que la population soit finie ou infinie.

- **Echantillonnage à partir d'une population finie (EASF)**

Un échantillon aléatoire simple de taille  $n$ , issu d'une population finie de taille  $N$ , est un échantillon sélectionné de manière à ce que chaque échantillon possible de taille  $n$  ait la même probabilité d'être sélectionné.

Il existe des procédures pour constituer un échantillon aléatoire simple à partir d'une population finie. Elle peut consister à effectuer des tirages successifs sans remise d'un élément dans un ensemble dénombrable et fini. De cette façon tous les éléments restant dans la population ont l'équiprobabilité d'être tirés.

Si les éléments tirés de la population-mère ne sont pas remis à leur place, la procédure correspond donc à la procédure d'échantillonnage sans remise. Par contre si les éléments déjà tiré, sont remis à leur place et peuvent de nouveau être tiré (possibilité qu'un élément soit tiré plusieurs fois) on est donc en présence d'une procédure d'échantillonnage avec remise.

NB : dans la pratique pour prélever un échantillon aléatoire simple, il faut :

- i. Constituer la base de sondage qui correspond à la liste complète et sans répétition des éléments de la population ;
- ii. Numéroté ces éléments de 1 à  $N$  ;
- iii. Procéder, à l'aide d'un générateur de nombres aléatoires (générés par ordinateur où se conférer à la table des nombres aléatoires) à la sélection des unités différentes qui constitueront l'échantillon aléatoire simple.

- **Echantillonnage à partir d'une population infinie (EASI)**

Dans certain cas, il peut arriver que la population finie, soit tellement importante qu'elle est traitée comme une population infinie pour des raisons pratiques. En pratique, une population étudiée est considérée comme infinie s'il est impossible de lister ou de compter tous les éléments de la population.

Un échantillon aléatoire simple, issu d'une population infinie, est un échantillon sélectionné de façon à ce que les conditions suivantes soient satisfaites :

- Chaque élément sélectionné provient de la même population
- Chaque élément est sélectionné de façon indépendante.

Dans ce cas une procédure de sélection d'échantillon doit être spécialement conçue pour sélectionner les éléments indépendamment et ainsi, éviter un biais de sélection qui procurerait à certains types d'éléments, des probabilités de sélection plus élevés.

- **Remarque**

Le nombre d'échantillons aléatoires simple différents de taille  $n$  qui peuvent être sélectionnés à partir d'une population de taille  $N$  est :

- Une combinaison de  $n$  dans  $N$ , si le tirage est fait sans remise
- Un arrangement avec répétition d'ordre  $n$  de  $N$ , si le tirage est fait avec remise.

**Exemple 41 :** Considérer une population finie composée de cinq éléments notés A, B, C, D et E. Dix échantillons aléatoires simples de taille égale à deux (2) peuvent être sélectionnés.

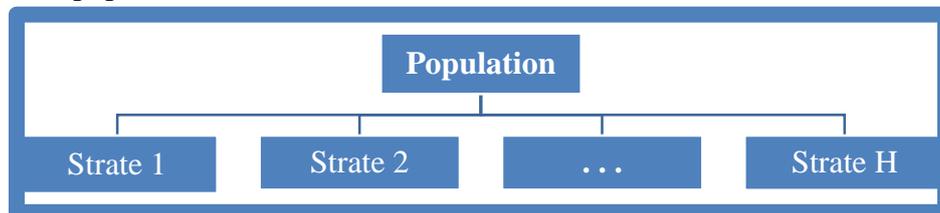
- Enumérer les dix échantillons en commençant avec AB, AC etc.
- En utilisant la procédure d'échantillonnage aléatoire simple, quelle est la probabilité pour chaque échantillon de taille deux d'être sélectionné ?

Supposez que le nombre aléatoire 1 correspond à A, le nombre aléatoire 2 correspond à B etc. Définir l'échantillon aléatoire de taille deux qui sera sélectionné en utilisant les chiffres 8 0 5 7 5 3 2.

### 2.1.2. Echantillonnage aléatoire stratifié

Dans l'échantillonnage aléatoire stratifié, une population hétérogène de taille  $N$  est tout d'abord divisé en  $p$  groupes d'éléments appelés strates, de façon à ce que  $N = N_1 + N_2 + \dots + N_p$  et chaque élément de la population appartienne à une et une seule strate ( $N_i$ ). L'élément de base qui définit une strate, tel qu'un service, un lieu, un âge, un type d'industrie etc. est laissé à la discrétion du créateur d'échantillon.

**Diagramme 1 :** population divisée en H strates



Après avoir formé les strates, un échantillon aléatoire simple d'effectif  $n_i$  est sélectionné dans chaque strate. Des formules permettent de combiner les résultats obtenus à partir des échantillons individuels en une estimation du paramètre de la population auquel on s'intéresse. Ainsi pour la répartition de l'effectif total  $n$ , de l'échantillon dans les différentes strates, la première solution, dite proportionnelle, consiste à conserver la même fraction d'échantillonnage dans chaque strate. Une seconde solution, dite optimale, tient compte du budget de l'enquête.

#### 2.1.2.1. Répartition proportionnelle

La répartition optimale consiste à répartir la taille de l'échantillon  $n$  en utilisant la même fraction de sondage  $f$  dans chacune des strates. Cette solution tient compte d'un seul facteur qui est le poids de chaque strate.

Désignons par  $w_i$  le poids de la strate et par  $f$  la fraction de sondage constante.

$$f = \frac{n}{N}; \quad w_i = \frac{N_i}{N}$$

Le nombre d'unités à choisir dans chacune des strates est donc :

$$n_i = w_i \times n = f \times N_i$$

**NB :**

- La stratification peut entraîner des gains de précision appréciables, elle facilite en outre les opérations de collecte des données et fournit des informations pour différentes parties de la population.
- Cette méthode fournit de meilleurs résultats lorsque la variance parmi les éléments de chaque strate est relativement faible.

**Exemple 42 : Répartition proportionnelle**

Dans une étude menée auprès d’une population de 10000 entreprises, réparties en 5000 Petites Entreprise, 3000 Moyennes entreprises et 2000 Grandes Entreprises, on souhaite avoir un échantillon de 500 entreprises.

Fraction de sondage constante :  $f = \frac{500}{10000} = 0.05\%$

Strate	Effectif de la strate	Taille de l'échantillon
Petite	5000	$5000 \times 0.05 = 250$
Moyenne	3000	$3000 \times 0.05 = 150$
Grande	2000	$2000 \times 0.05 = 100$
TOTAL	10000	500

**2.1.2.2. Répartition optimale**

Cette deuxième solution consiste à répartir l’effort d’échantillonnage de façon inégale dans les différentes strates. Elle tient compte de quatre facteurs :

- Le budget total de l’enquête, G
- Le poids de la strate,  $w_i$
- Le coût de la collecte de l’information dans la strate,  $c_i$
- La dispersion à l’intérieur de la strate, mesurée par l’écart type  $\sigma_i$ .

Le nombre d’unités à choisir dans chacune des strates est donné par :

$$n_i = k \frac{w_i \sigma_i}{\sqrt{c_i}} \text{ avec } k = \frac{G}{\sum w_i \sigma_i \sqrt{c_i}}$$

**Exemple 43:** Dans une population des 10000 entreprises, on a pu avoir les informations suivantes :

Strate	Poids de la strate $w_i$	Coût de la collecte de l’information dans la strate $c_i$	Dispersion à l’intérieur de la strate, mesurée par l’écart type $\sigma_i$
Petite	0.5	50	0.8
Moyenne	0.3	75	1.5
Grande	0.2	100	2.2

le nombre d’entreprises à choisir dans chacune des strates est donné par :

$$k = \frac{G}{\sum w_i \sigma_i \sqrt{c_i}} = \frac{5000}{0.5 \times 0.8 \times \sqrt{50} + 0.3 \times 1.5 \times \sqrt{75} + 0.2 \times 2.2 \times \sqrt{100}} = 449,42$$

$$n_i = k \frac{w_i \sigma_i}{\sqrt{c_i}}$$

$$n_1 = 449.42 \times \frac{0.5 \times 0.8}{\sqrt{50}} = 26 \text{ Petites entreprises}$$

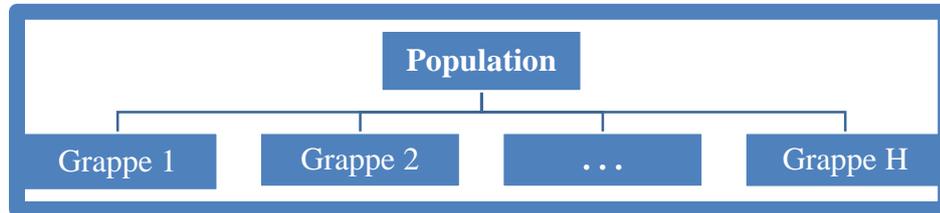
$$n_2 = 449.42 \times \frac{0.3 \times 1.5}{\sqrt{75}} = 24 \text{ Moyennes entreprises}$$

$$n_3 = 449.42 \times \frac{0.2 \times 2.2}{\sqrt{100}} = 20 \text{ Grandes entreprises}$$

### 2.1.3. Echantillonnage par grappes

Dans l'échantillonnage par grappe, la population est tout d'abord divisé en groupes d'éléments séparés appelés *grappes*. Un échantillon aléatoire simple des grappes est ensuite sélectionné. Tous les éléments contenus dans une grappe sélectionnée forment l'échantillon.

**Diagramme 2 :** population divisée en H grappes



**Exemple 44:** Echantillonnage de région, ou les grappes sont les quartiers d'une ville ou d'autres régions bien définies.

**NB :**

- contrairement à l'échantillonnage aléatoire stratifié, cette méthode fournit de meilleurs résultats lorsque la variance parmi les éléments de chaque strate est considérable. C'est-à-dire lorsque les éléments dans la grappe sont hétérogènes (dissemblable);
- un échantillonnage par grappe nécessite généralement un échantillon total de taille plus importante que l'échantillonnage aléatoire simple ou stratifié. Par conséquent elle permet de faire des économies considérables.

### 2.1.4. Echantillonnage systématique

Lorsque les populations sont importantes, il est coûteux (en temps) de sélectionner un échantillon aléatoire simple. L'échantillonnage systématique est une méthode d'échantillonnage probabiliste dans laquelle on choisit aléatoirement un des  $k$  premiers éléments, puis tous les  $k^{\text{e}}$  éléments qui suivent. Cette méthode constitue une alternative à l'échantillonnage aléatoire simple.

De manière pratique, si on connaît l'effectif total de la population  $N$  et qu'on souhaite prélever un échantillon d'effectif  $n$ , l'intervalle entre deux unités successives à sélectionner est donné par :  $k = \frac{N}{n}$  (Arrondi à l'entier le plus proche)

Connaissant  $k$ , on choisit le plus souvent, pour débiter, un nombre aléatoire,  $i$ , compris entre 1 et  $k$ . le rang des unités sélectionnées est alors  $i, i + 2k, i + 3k, \dots$

L'échantillonnage systématique est facile à préparer et, en général facile à exécuter, il 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 et simple. Par contre Si les éléments de la population présentent une tendance, l'échantillonnage systématique est plus précis que l'échantillonnage aléatoire.

Exemple 45: soit à sélectionner un échantillon de taille 50 parmi une population contenant 5000 éléments. Cela revient à sélectionner un élément tous les  $k = 5000/50 = 100$  éléments de la population.

Constituer un échantillon systématique dans ce cas revient à tirer un élément de la population tous les 100 en partant d'un nombre tiré au hasard entre 1 à 100.

Supposons que ce nombre est le 27. On va donc sélectionner le 27<sup>ème</sup> élément puis le 127<sup>ème</sup>, le 227<sup>ème</sup>, le 327<sup>ème</sup>, ..., 10027<sup>ème</sup> élément ce qui nous donnera l'échantillon de 100 éléments.

## 2.2. Méthodes d'échantillonnage non-probabiliste

Cette section fournit une brève introduction aux méthodes d'échantillonnage autres que la méthode d'échantillonnage probabiliste. Ces méthodes sont dites « empiriques ». L'échantillonnage par commodité et l'échantillonnage subjectif, sont les deux méthodes empiriques le plus utilisés et seront donc étudiés dans cette section.

### 2.2.1. Échantillonnage de commodité (ou convenance)

Comme son nom l'indique, l'échantillon est principalement identifié par commodité c'est-à-dire en fonction de leur disponibilité ou facilité. Les éléments sont donc inclus dans l'échantillon sans probabilité connue ou pré-spécifiée d'être choisis. L'échantillonnage par commodité se définit donc comme une méthode non-probabiliste dans laquelle les éléments de l'échantillon sont sélectionnés en fonction de leur commodité.

Exemple 46: un professeur qui mène une expérience à l'université peut utiliser des étudiants volontaires pour constituer un échantillon simplement parce qu'ils sont disponibles et participeront en tant que sujet à l'expérience pour un coût très faible ou même nul.

**NB :**

Cette méthode ne donne pas les mêmes chances à tous les échantillons possibles d'être sélectionnés ; ainsi elle induit un risque d'erreur non contrôlé.

si un échantillon de commodité est facile à être constituer son principal défaut est qu'aucune procédure statique bien fondée ne permet de faire une analyse probabiliste ou de l'inférence sur la qualité des résultats de l'échantillon. Toutefois il arrive des fois où des chercheurs appliquent à des échantillons de commodité des méthodes statistiques, cela n'a aucun fondement théorique solide.

### 2.2.2. Échantillonnages subjectifs

Dans cette approche la personne la mieux documentée sur le sujet de l'étude sélectionne des éléments de la population qu'elle pense être les plus représentatif de la population. En d'autres termes, un échantillonnage subjectif est une méthode d'échantillonnage non probabiliste dans laquelle les éléments de l'échantillon sont sélectionnés en fonction des croyances de la personne qui fait l'étude.

**Sondage par la méthode des quotas :** lorsque le chercheur veut reproduire les caractéristiques d'une population tel que l'âge, le sexe, les revenus...

**Échantillonnage par jugement :** lorsque le chercheur juge que l'échantillon va lui permettre d'atteindre les objectifs de la recherche,

**Échantillonnage boule de neige :** Utile dans le cas de la rareté des unités d'échantillonnage ou de l'absence d'un cadre d'échantillonnage valide. On demande à un répondant de nous référer à un autre qui présente les mêmes caractéristiques que les siennes, et ainsi de suite...

**NB :** Cette méthode ne donne pas les mêmes chances à tous les échantillons possibles d'être sélectionnés ; ainsi elle induit un risque d'erreur non contrôlé.

### III. DISTRIBUTION D'ÉCHANTILLONNAGES

#### 3.1. Position du problème

Si nous prélevons un échantillon de taille  $n$  dans une population donnée, la moyenne de l'échantillon nous donnera une idée approximative de la moyenne de la population. Seulement si nous prélevons un autre échantillon de même taille, nous obtiendrons une autre moyenne d'échantillon. Sur l'ensemble des échantillons possibles, on constatera que certains ont une moyenne proche de la moyenne de la population et que d'autres ont une moyenne qui s'en écarte davantage. Comment traiter le problème ? Un échantillon de taille  $n$  (appelé aussi un  $n$ -échantillon), obtenu par échantillonnage aléatoire, va être considéré comme le résultat d'une expérience aléatoire. A chaque échantillon de taille  $n$  on peut associer la valeur moyenne des éléments de l'échantillon. On a donc défini une variable aléatoire qui à chaque  $n$ -échantillon associe sa moyenne échantillonnage. On la note  $\bar{X}_i$ . Cette variable aléatoire possède bien entendu :

- Une distribution de probabilité.
- Une valeur moyenne (la moyenne des moyennes d'échantillons).
- Un écart-type.

Le but de ce paragraphe est de déterminer ces trois éléments. Avant de continuer, essayons de comprendre sur un exemple ce qui se passe.

#### 3.2. Exemple introductif

Pour mettre en place une nouvelle politique de bourse, le nouveau gouvernement centrafricain a chargé la Faculté des Sciences Economiques et de Gestion (FASEG) d'identifier le profil des **21.500** étudiants que compte l'Université de Bangui (UB). Les caractéristiques pertinentes comprennent le revenu moyen annuel et la proportion des étudiants ayant déjà bénéficié de la bourse du gouvernement.

En considérant les **21500** étudiants comme population de cette étude, on peut déterminer le revenu annuel moyen de chaque individu, et savoir s'il a déjà bénéficié d'une bourse du gouvernement.

En utilisant l'ensemble des données de l'UB et les formules vues en statistique descriptive, nous pouvons calculer la moyenne et l'écart-type du revenu annuel moyen pour la population :

- Moyenne de la population :  $m = 51800 \text{ FCFA}$
- Ecart-type de la population :  $\sigma = 4000 \text{ FCFA}$

Les données concernant le programme de bourse montre que 13000 des 21500 étudiants ont une fois bénéficié d'une bourse du gouvernement. Soit  $p$  la proportion des étudiants ayant bénéficié d'une bourse du gouvernement. Nous avons donc  $p = \frac{13000}{21500} = 0.60$

Le revenu annuel moyen ( $m = 51800 \text{ FCFA}$ ), l'écart-type du revenu annuel moyen de la population ( $\sigma = 4000 \text{ FCFA}$ ) et la proportion des étudiants ayant bénéficié d'une bourse du gouvernement. ( $p = 0.60$ ) sont les paramètres de la population des étudiants de l'université de Bangui.

### 3.3. Distribution d'échantillonnage

Considérons tous les échantillons de taille  $n$  qui peuvent être extraits au hasard (avec ou sans remise) de cette population (finie ou infinie).

Chaque échantillon constitue une distribution statistique, pour laquelle on peut calculer une caractéristique de la distribution statistique, telle que *la moyenne, l'écart-type, le mode, une proportion...* cette caractéristique va varier d'un échantillon à l'autre. Ainsi pour la totalité des échantillons prélevés dans la population, on peut obtenir plusieurs valeurs de la caractéristique étudiée (*moyenne, l'écart-type, le mode, une proportion ou ...*) on dit alors que l'on est en présence d'une *distribution d'échantillonnage* concernant la caractéristique étudiée.

**NB :** on parle d'une *distribution d'échantillonnage de la moyenne* lorsqu'on est dans le cas de la *moyenne* ; d'une *distribution d'échantillonnage de proportion* dans le cas de *proportion...*

Suivant le caractère de la population (finie ou infinie), suivant le type de tire (avec ou sans remise), il existe des relations entre les caractéristiques de la distribution d'échantillonnage et les caractéristiques des échantillons.

#### 3.3.1. Notation

Les difficultés constatées dans l'étude des distributions d'échantillonnage proviennent souvent d'une confusion dans les notations. Il est indispensable de séparer, du point de vue notation comme du point de vue raisonnement, ce qui appartient à la population-mère, de ce qui appartient aux échantillons, et ce qui appartient aux échantillons, de ce qui appartient à la distribution d'échantillonnage.

Soit une population de  $N$  sujets. A chaque sujet est attachée une caractéristique  $x_i$ . On désigne la moyenne du caractère  $x_i$ , dans la population, par  $m$  :

$$m = \frac{\sum_{i=1}^N x_i}{N}$$

L'écart-type du caractère  $x_i$  dans la population, par  $\sigma$  :

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - m)^2}{N}}$$

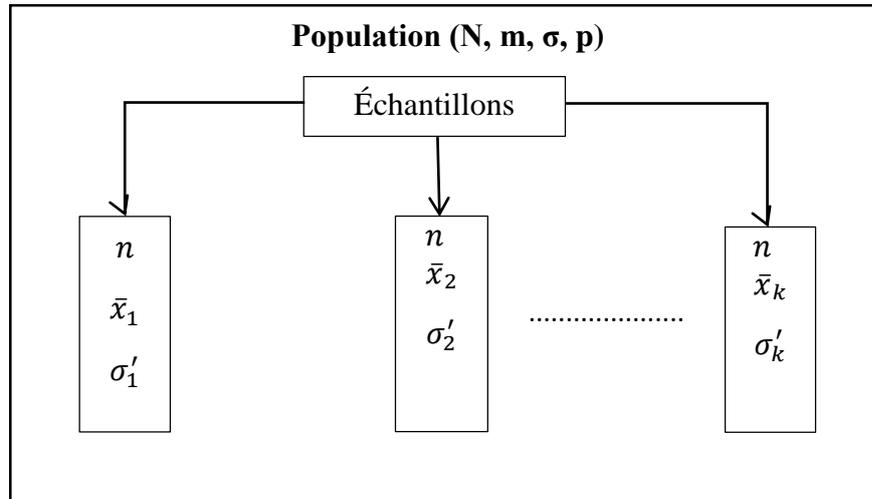
On prélève à l'intérieur de la population mère,  $k$  échantillons de taille  $n$  et on désigne par  $\bar{x}_1$  la moyenne et par  $\sigma'_1$  l'écart-type du caractère  $x_i$  dans le premier échantillon ( $E_1$ ), par  $\bar{x}_2$  la moyenne et par  $\sigma'_2$  l'écart-type du caractère  $x_2$  dans le deuxième échantillon ( $E_2$ ), ... par  $\bar{x}_k$  la moyenne et par  $\sigma'_k$  l'écart-type du caractère  $x_k$  dans le  $k^{\text{ième}}$  échantillon ( $E_k$ ).

L'ensemble  $\bar{X} = \{\bar{x}_1, \bar{x}_2, \bar{x}_3 \dots \bar{x}_k\}$  d'effectif  $k$  constitue une *distribution d'échantillonnage de moyennes*, notée  $\bar{X}$

Souvent on désigne par  $\mu$ , la moyenne et  $\sigma_{\bar{x}}$  l'écart-type de la *distribution d'échantillonnage (de moyenne)*

### 3.3.2. Illustration graphique

Il est commode de visualiser ces différentes notations sur un schéma :



$\bar{X}$ : Distribution d'échantillonnage ( $\mu, \sigma_{\bar{X}}$ )

Dans le cas d'une distribution d'échantillonnage de proportion, on remplace usuellement, du point de vue notation, le caractère  $x_i$  par la lettre  $f$ .

### 3.3.3. Interprétation probabiliste

Nous avons définis une variable aléatoire comme une application qui permet d'associer à une expérience aléatoire une valeur numérique. Ainsi prélever au hasard avec une équiprobabilité, dans une population-mère P (une expérience aléatoire) un sujet auquel est attaché un caractère  $x_i$  (valeur numérique) revient à définir une variable aléatoire  $X$ . la moyenne de cette échantillon  $\bar{x}$  est aussi une variable aléatoire. Par conséquent, comme pour toute variable aléatoire,  $\bar{X}$  a :

- une espérance mathématique  $E(\bar{X}) = m$
- une variance  $V(\bar{X}) = \sigma_{\bar{X}}^2$
- et une distribution de probabilité

La distribution de probabilité de  $\bar{X}$  est appelée **distribution d'échantillonnage de  $\bar{X}$**

### 3.3.4. distribution d'échantillonnage de $\bar{X}$

Nous venons de montrer que la moyenne d'échantillon  $\bar{X}$  est une variable aléatoire et sa distribution de probabilité est appelée **distribution d'échantillonnage de  $\bar{X}$** .

La distribution d'échantillonnage de  $\bar{X}$  correspond à la distribution de probabilité de toutes les valeurs possibles de la moyenne d'échantillon  $\bar{X}$ . Ainsi comme tout autres distribution de probabilité, la distribution d'échantillonnage de  $\bar{X}$  a une espérance mathématique, un écart-type et une forme caractéristique.

#### ✚ Espérance mathématique de $\bar{X}$

Soit  $E(\bar{X})$  l'espérance mathématique de  $\bar{X}$  et  $m$  la moyenne de la population d'où est issu un échantillon aléatoire simple. On peut montrer qu'avec un EAS  $E(\bar{X})$  et  $\mu$  sont égaux.

$$E(\bar{x}) = m$$

Où

$E(\bar{x})$  Correspond à l'espérance mathématique

$m$  Correspond à la moyenne de la population.

L'espérance mathématique de  $\bar{X}$  est égale à la *moyenne de la population* d'où est issu l'échantillon.

**✚ Ecart-type**

Définissons l'écart-type de la distribution d'échantillonnage de  $\bar{X}$ . Nous utiliserons les notations suivantes :

$\sigma_{\bar{x}}$  pour l'écart-type de la distribution d'échantillonnage de  $\bar{X}$ .

$\sigma$  pour l'écart-type de la population

$n$  pour la taille de l'échantillon

$N$  pour la taille de la population

**✚ Pour une population finie**

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \left( \frac{\sigma}{\sqrt{n}} \right)$$

**✚ Pour une population infinie**

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

En comparant ces deux expressions ; il apparait que le facteur  $\sqrt{\frac{N-n}{N-1}}$  est nécessaire lorsque la population est finie, mais pas lorsqu'elle est infinie. Ce facteur est généralement appelé *facteur de correction pour une population finie*. Dans de nombreux échantillonnage la population bien que finie ; est « grande », alors que la taille de l'échantillon est relativement « petite » dans de tel cas le facteur de correction est petite.

**✚ Remarque :**

i.  $\lim_{N \rightarrow +\infty} \sqrt{\frac{N-n}{N-1}} = 1$

Par conséquent  $\frac{\sigma}{\sqrt{n}}$  devient une bonne approximation de l'écart-type  $\sigma_{\bar{x}}$  même si la population est finie. Du point de vue empirique, il suffit que  $\frac{n}{N} \geq 0.05$ . C'est à dire que l'échantillon soit inférieur ou égal à 5% de la taille de la population finie.

ii. Le calcul de l'écart-type d'un échantillon  $\sigma_{\bar{x}}$  nécessite la connaissance de l'écart-type de la population  $\sigma$ . C'est pourquoi l'écart-type d'un échantillon  $\sigma_{\bar{x}}$  est souvent appelé **erreur type** de la moyenne.

**Exemple 47:** Dans l'exemple introductif l'écart-type du revenu annuel moyen de la population est égal à 4000FCFA. La population est finie  $N = 21500$  FCFA, cependant la taille de l'échantillon  $n = 30$ , nous avons donc  $n/N = 30/21500 = 0.001$  donc inférieur à 5%. Ce qui nous permet d'écrire :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4000}{\sqrt{30}} = 730.3$$



### 3.3.5. Forme de la distribution d'échantillonnage de $\bar{x}$

La détermination de la distribution d'échantillonnage de  $\bar{x}$  permet de considérer deux cas. Soit la population à une distribution normale (c'est le cas idéal), soit elle n'a pas de distribution normale. Dans ce cas le théorème central limite permet de résoudre le problème.

#### ✚ Théorème central limite

En sélectionnant des échantillons aléatoires simples de taille  $n$  d'une population, la distribution d'échantillonnage de la moyenne d'échantillon  $\bar{x}$  peut-être approchée par une distribution de probabilité normale lorsque la taille de l'échantillon devient importante.

**NB :** dans la pratique, pour  $n \geq 30$  :  $L(\bar{x}) \sim N(\mu, \sigma_{\bar{x}})$  si  $N > 2n$

### 3.2. distribution d'échantillonnage de $\bar{p}$

Soit une population mère composée de  $N$  sujets. Chaque sujet possède ou ne possède pas un caractère. La proportion de ceux qui le possède est  $p$  ; la proportion de ceux qui ne le possède pas est  $q = 1 - p$ .

Considérons tous les échantillons de taille  $n$  qui peuvent être extraits de la population. Pour chaque échantillon déterminons la proportion  $f$  de succès. Nous obtenons une distribution d'échantillonnage de proportion  $F$ . soit  $p$  la moyenne de  $F$  et  $\sigma_F$  l'écart-type de  $F$

#### a) Espérance mathématique de $\bar{p}$

Soit  $E(\bar{p})$  l'espérance mathématique de  $\bar{p}$  et  $\mu$  la moyenne de la population d'où est issu un échantillon aléatoire simple. On peut montrer qu'avec un EAS  $E(\bar{p})$  et  $p$  sont égaux.

$$E(\bar{p}) = p$$

Où

$E(\bar{p})$  Correspond à l'espérance mathématique de  $\bar{p}$

$p$  Correspond à la proportion de la moyenne.

Puisque  $E(\bar{p}) = p$ ,  $\bar{p}$  est un estimateur sans biais de  $p$

#### b) Ecart-type de $\bar{p}$

Comme pour la moyenne d'échantillon  $\bar{x}$ , l'écart-type de  $\bar{p}$  est différent la population est finie ou infinie:

#### ✚ Pour une population finie

$$\sigma_{\bar{p}} = \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{p(1-p)}{n}}$$

#### ✚ Pour une population infinie

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

**Exemple 48:** Dans l'exemple introductif nous savons que la proportion des étudiants ayant bénéficié d'une bourse du gouvernement est égal à 0.60 avec  $n/N = 30/21500 = 0.001$ , nous pouvons ignorer le facteur d'exhaustivité et calculer l'erreur type de la proportion.

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.6(1-0.6)}{30}} = 0.0894$$

En considérant les caractéristiques des distributions d'échantillonnage de  $\bar{x}$ , et  $\bar{p}$ , sous certaines conditions (comme nous l'avons vu dans le chapitre 3) elles peuvent suivre une loi normale.

### .3.3. Application du théorème central limite

Pour de grand échantillon ( $n \geq 30$ ) la loi de F est approximativement distribuée suivant une loi normale bien que la population suit une distribution de loi binomiale.

$$\text{Pour } n \geq 30 : L(F) \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

### .3.4. Distribution spécifiée

Si la variable aléatoire X, dont les réalisations  $x_i$  sont associées aux sujets de la population-mère, obéit à une loi de probabilité particulière, la distribution d'échantillonnage est influencée par cette loi. Par exemple, si X obéit à une loi normale, même si  $n$  est inférieur à 30.  $L(F) = \mathcal{N}(m, \sigma) \Rightarrow L(\bar{X}) = \mathcal{N}(\mu, \sigma_{\bar{X}})$

# C

## HAPITRE V

### ESTIMATION

A partir des données de la population-mères, la théorie d'échantillonnage permet de déduire des résultats au sujet des échantillons extraits de la population. Le problème de l'estimation est le problème inverse. En effet, lorsque les paramètres d'une population sont inconnus on peut s'attacher à les estimer à partir des valeurs (moyenne, écart-type, ...) d'échantillons représentatifs extraits de cette population-mère.

Du point de vue utilitaire, ce dernier problème est plus important que le problème contraire, car devant la difficulté de recourir à des *recensements*, le seul moyen dont dispose l'économiste ou le gestionnaire pour connaître les paramètres (moyenne, écart-type, ...) d'une population réside en l'estimation. Pour estimer les paramètres d'une population à partir d'échantillons représentatifs, on peut adopter deux attitudes (qui ne sont d'ailleurs pas exclusives l'une de l'autre) : ou bien on cherche à attribuer au paramètre inconnu une valeur, la plus raisonnable possible compte tenu des observations, ou bien on cherche à le situer dans un intervalle ayant une probabilité donnée de recouvrir la valeur inconnue. La valeur unique choisie s'appelle *estimation ponctuelle*, alors que l'intervalle s'appelle *intervalle de confiance* du paramètre inconnu. Ces deux modes d'estimation, feront respectivement l'objet des deux sections de ce chapitre.

#### I. ESTIMATION PONCTUELLE D'UN PARAMETRE

Estimer un paramètre  $\theta$  inconnu par un nombre constitue une estimation ponctuelle du paramètre inconnu.

##### 1.1. Définition

Soit un caractère  $X$  (une variable aléatoire) d'une population quelconque, dont la loi de probabilité  $L(X)$  a pour fonction de densité de probabilité  $f(x, \theta)$ , laquelle dépend d'un paramètre inconnu  $\theta$  à estimer. Soient  $x_1, x_2, \dots, x_n$  les valeurs prises par  $X$  dans un échantillon de taille  $n$ .

On appelle estimateur  $T_n$  du paramètre  $\theta$ , la fonction qui aux valeurs de l'échantillon fait correspondre la valeur du paramètre  $\theta$ .

$$f: X \rightarrow \Theta$$

$$(x_1, x_2, \dots, x_n) \mapsto T_n = f(x_1, x_2, \dots, x_n)$$

La fonction  $T_n$  est une fonction numérique d'un échantillon aléatoire. C'est donc une variable aléatoire.

On appelle *estimations* les valeurs numériques  $t_1, t_2, \dots, t_n$  de cette variable aléatoire  $T_n$ .

Estimer le paramètre  $\theta$  consiste à donner une valeur approchée à ce paramètre à partir d'un sondage de la population.

**Exemple 49 :** Soit  $X$  une variable aléatoire et  $X_1, X_2, \dots, X_n$  un échantillon de taille  $n$ . Alors

$$\hat{\theta}_1 = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$\hat{\theta}_1 = X_i$  (Une valeur individuelle) sont des estimateurs de la moyenne  $\mu = E(X)$ .

**Exemple 50 :** Soit  $X$  une variable aléatoire et  $X_1, X_2, \dots, X_n$  un échantillon de taille  $n$ . Alors

$$\hat{\theta}_1 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$\hat{\theta}_1 = S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  sont des estimateurs de la variance  $\sigma^2 = V(X)$ .

### 1.2. Propriétés d'un estimateur

Pour qu'un estimateur soit valable on exige de lui certaines qualités. De ces qualités les plus essentielles sont, le biais, la convergence et l'optimalité.

#### 1.2.1. Biais d'un Estimateur

Pour pouvoir considérer  $T_n = (x_1, x_2, \dots, x_n)$  comme une valeur approchée de  $\theta$ , il faut que les valeurs prises par la variable aléatoire  $T_n$  ne s'écartent pas trop de la valeur fixe de  $\theta$ . Comme  $T_n$  est une v.a, on ne peut imposer une condition qu'à sa valeur moyenne, ce qui nous amène à définir le biais d'un estimateur comme l'écart entre sa moyenne et la vraie valeur du paramètre :

$$B_n(\theta) = E_\theta(T_n) - \theta$$

##### a) Estimateur sans biais

Un estimateur  $T_n$  de  $\theta$  est dit sans biais si l'espérance mathématique de l'estimateur est égale à la vraie valeur du paramètre  $\theta$  :

$$E_\theta(T_n) = \theta$$

$$B_n(\theta) = E_\theta(T_n) - \theta$$

##### b) Estimateur asymptotiquement sans biais

Un estimateur  $T_n$  de  $\theta$  est dit asymptotiquement sans biais si

$$\lim_{n \rightarrow \infty} E_\theta(T_n) = \theta$$

**Exemple 51 :** Soit  $X$  une variable aléatoire et  $X_1, X_2, \dots, X_n$  un échantillon de taille  $n$ . Alors

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^5 X_i \text{ et } \hat{\theta}_1 = X_i = \frac{2X_1 - X_2 + X_4}{2}$$

1. Ces deux estimateurs sont-ils non-biaisés ?
2. Quel est le meilleur des deux ?

#### 1.2.2. Convergence d'un estimateur

Un estimateur  $T_n$  est convergent si pour tout  $\epsilon > 0$  la suite de variables aléatoire ( $T_n$ ) converge en probabilité vers la valeur du paramètre :

$$\lim_{n \rightarrow \infty} P(|T_n - \theta| < \epsilon) = 1$$

Cela signifie que si la taille de l'échantillon est assez grande alors on est (presque) certain que l'estimateur  $T_n$  est très proche de  $\theta$

- **Théorème**

Tout estimateur sans biais ou asymptotiquement sans biais, dont la variance tend vers zéro, quand  $n$  tend vers l'infini, est convergent.

### 1.2.3. Estimateur optimal

#### 1.2.3.1. mesure de la précision d'un estimateur : l'Erreur Quadratique Moyenne.

La précision d'un estimateur se mesure par l'*erreur quadratique moyenne*, définie pour tout  $\theta$  par :

$$EQM(T_n) = E_{\theta}(T_n - \theta)^2 = V_{\theta}(T_n) + B_{\theta}^2(\theta)$$

Dans le cas particulier d'un estimateur sans biais on a :

$$EQM(T_n) = E_{\theta}(T_n - \theta)^2 = V_{\theta}(T_n) \text{ avec } B_{\theta}^2(\theta) = 0$$

Un bon estimateur ponctuel doit être précis : il l'est d'autant plus que son erreur quadratique est faible (biais et variance les plus faibles possibles)

Cette erreur quadratique se confond donc avec la variance de l'estimateur. Ainsi, si dans l'erreur total d'estimation on privilégie l'erreur structurel, mesurée par  $B_{\theta}^2(\theta)$ , on fera le choix d'un estimateur sans biais et l'erreur d'estimation se réduira à l'erreur statistique mesuré par la variance de l'estimateur.

**Théorème :** si  $EQM(T_n)$  converge vers 0 lorsque  $n$  tend vers infini alors  $T_n$  est convergent.

#### 1.2.3.2. Estimateur à variance minimal

Un estimateur de  $\theta$  est dit de variance minimale si, parmi tous les estimateurs possibles de  $\theta$  il a la plus petite variance :

Considérant qu'on se place dorénavant dans ce cours dans la classe des estimateurs sans biais, on pourra comparer deux estimateurs  $T_n$  et  $T_n'$  de même classe par leur variance, qui mesure alors leur dispersion par rapport au paramètre, qui est leur espérance commune. Nous dirons que l'estimateur  $T_n$  est plus efficace que  $T_n'$  si pour tout  $\theta$  et pour une taille d'échantillon  $n > N$  :

$$V_{\theta}(T_n) \leq V_{\theta}(T_n')$$

**Théorème :** Si nous considérons tous les estimateurs possibles dont les distributions d'échantillonnage ont la même moyenne, l'estimateur le meilleur est celui dont la variance est minimum, à effectif égal de l'échantillon.

## 1.3. Inégalité de Cramer-Rao

### 1.3.1. Définition

On appelle vraisemblance (likelihood) de l'échantillon  $(x_1, x_2, \dots, x_n)$  la loi de probabilité de ce n-uple, notée  $L(x_1, x_2, \dots, x_n; \theta)$  et définie par :

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n P(X_i = x_i | \theta)$$

Si  $x$  est une variable aléatoire discrète, et par :

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

Si  $x$  est une variable aléatoire continue de densité de probabilité  $f(x_i; \theta)$

**1.3.2. Information de Fisher d'un échantillon ( $\theta$  réel)**

Score de l'échantillon : Si  $f(x, \theta)$  est différentiable en  $\theta$ ,  $L$  est une fonction dérivable de  $\theta$ , et le score est sa dérivée :  $S_n(\theta) = \frac{\partial}{\partial \theta} L(x_1, x_2, \dots, x_n; \theta)$

Pour tout  $\theta$ , Le score est une variable aléatoire. Le score s'annule à un optimum en  $\theta$  de la fonction de vraisemblance. Le score est centré :

$$E(S_n(\theta)) = 0$$

La variance du score (si elle existe) s'appelle *l'information de Fisher apportée par l'échantillon sur  $\theta$*  :

$$I_n(\theta) = E((S_n(\theta))^2)$$

$$I_n(\theta) = E_{\theta} \left( \frac{\partial^2 L}{\partial \theta^2} \right)$$

Cette quantité mesure l'information apportée par un échantillon sur le paramètre

**1.3.3 Propriétés de l'information de Fisher :**

1. Si le domaine de définition de  $X$  ne dépend pas de  $\theta$  et que cette quantité existe,

$$I_n(\theta) = -E_{\theta} \left( \frac{\partial^2 L}{\partial \theta^2} \right) = -E_{\theta} \left( \frac{\partial S_n(\theta)}{\partial \theta} \right)$$

2. Additivité : si le domaine de définition de  $X$  ne dépend pas de  $\theta$ , chaque observation apporte la même information :

$$I_n(\theta) = nI_1(\theta)$$

**1.3.4. Théorème :**

Sous les hypothèses de Cramer-Rao, en particulier si  $E = X(\Omega)$  est indépendant du paramètre à estimer  $\theta$ , pour tout estimateur sans biais  $T_n$  de  $\theta$  on a :

$$V_{\theta}(T_n) \geq \frac{1}{I_n(\theta)} = b_F(\theta)$$

Où

- $I_n(\theta)$  est la *quantité d'information (information de Fisher)* :

$$\bullet \quad I_n(\theta) = E_{\theta} \left[ \left( \frac{\partial L}{\partial \theta} \right)^2 \right]$$

- Notons que dans les conditions d'application de ce théorème, en particulier si  $E = X(\Omega)$  est indépendant du paramètre à estimer on obtient une expression équivalente de la qualité d'information de Fischer, qui est généralement plus simple à calculer :

$$I_n(\theta) = nE_{\theta} \left( \frac{\partial^2 L}{\partial \theta^2} \right)$$

Cette quantité mesure l'information apportée par un échantillon sur le paramètre : une information de Fisher proche de zéro indique un échantillon peu informatif sur la valeur de  $\theta$ .

- $b_F(\theta)$  est la borne inférieure de **Fréchet-Darmoi-Cramer-Rao** (FDCR en abrégé).

**1.4. Estimateur efficace**

Un estimateur sans biais  $T_n$  est dit efficace si la variance de l'estimateur est égale au minimum (borne inférieure) fournie par l'inégalité FDCR :

$$V_{\theta}(T_n) = \frac{1}{I_n(\theta)}$$

**Interprétation :** La valeur de la borne de FDCR est fonction de l'information que peut contenir l'échantillon sur le paramètre : Plus grande est l'information sur la valeur du paramètre, plus précises seront les prédictions d'un estimateur sans biais dont la variance est égale à la borne de Cramer-Rao. Inversement, si la variance du score (information de Fisher) est très petite, et donc si presque tous les échantillons ne contiennent que peu d'information sur la valeur du paramètre, on ne peut pas espérer d'un estimateur sans biais qu'il soit précis, c'est à dire qu'il ait une faible variance.

**Remarque :** L'inégalité de FDCR et donc la variance minimale que peut atteindre un estimateur sans biais, n'est valable que dans le cas où l'information de Fisher existe et où H est vérifiée.

- **Remarque :**

1. La propriété la plus désirable pour un estimateur est d'avoir une faible Erreur Quadratique Moyenne (ce qui n'exige pas forcément d'être sans biais), mais la difficulté est que la théorie de l'estimation ne permet pas de résoudre le problème de minimisation de l'EQM (fonction dépendant de manière complexe du paramètre).

Un compromis consiste donc à rechercher un estimateur sans biais de variance minimale :

2. L'absence de biais facilite grandement l'étude des propriétés d'un estimateur car le biais d'un estimateur peut dépendre de façon complexe de la valeur du paramètre.
3. Il est cependant possible de trouver des estimateurs biaisés plus précis que le meilleur estimateur sans biais.

Ce compromis soulève un autre problème qui est celui du calcul de la variance d'un estimateur. En effet, l'existence et la définition d'un estimateur de variance minimale nécessite généralement la connaissance de la loi de probabilité jointe de l'échantillon aléatoire, ce qui n'est pas possible. Car comme on l'a montré ci-haut, la fonction de densité de cette loi de probabilité dépend à son tour d'un paramètre inconnu.

2. L'absence de biais ne garantit pas la plus faible valeur possible de l'EQM : celle-ci sera atteinte lorsque sera trouvé le meilleur compromis entre le biais de l'estimateur et sa variance entre le biais de l'estimateur et sa variance.
3. Dans certains cas même, l'introduction d'un léger biais dans un estimateur initialement sans biais peut conduire à une réduction significative de sa variance, au point de provoquer une diminution de son EQM, et donc d'améliorer ses performances.

## II. METHODES DE CONSTRUCTION D'UN ESTIMATEUR

### 2.1.Méthode du Maximum de Vraisemblance (MMV)

Le principe de cette méthode est de choisir comme estimation de tout paramètre  $\theta$  la valeur la plus vraisemblable, c'est à dire celle qui a la plus grande probabilité de provoquer l'apparition des valeurs observées dans l'échantillon. Cette probabilité est appelée fonction de vraisemblance.

#### 2.1.1. Définition

On appelle Estimateur du Maximum de Vraisemblance (EMV) toute fonction  $\hat{\theta}_n$  de  $(x_1, x_2, \dots, x_n)$  qui vérifie :

$$L(x_1, x_2, \dots, x_n; \hat{\theta}_n) = \max_{\theta_i \in \Theta} L(x_1, x_2, \dots, x_n; \theta)$$

**2.1.2. Mise en œuvre**

Cette définition ne renseigne en rien ni sur l'existence, ni sur l'unicité d'un tel estimateur. Dans la plupart des cas, on maximisera la log-vraisemblance en annulant sa dérivée par rapport à chaque composante de  $\theta$  (i. *L'Approche standard*). Mais cette méthode ne fonctionnait pas toujours. C'est pourquoi il faut parfois recourir à d'autres approches (la méthode (ii. *log-vraisemblance concentré* ou iii. la méthode *log-vraisemblance conditionnel*).

Ainsi nous voyons que la mise en œuvre de la MMV peut se faire de trois manières :

- (a) L'Approche *standard*
- (b) la méthode *log-vraisemblance concentré*.
- (c) la méthode *log-vraisemblance conditionnel*.

**a) L'Approche standard**

Elle comporte deux étapes :

**Etape 1 :** spécification de la fonction de *vraisemblance (log-vraisemblance)*

**Etape 2 :** maximisation de la fonction de *vraisemblance (log-vraisemblance)*

- ✓ Identifier les extrema de la *vraisemblance (log-vraisemblance)* Retenir parmi ces extrema ceux qui sont des maxima:

$$\frac{dL}{d\theta} = 0; \text{ ou } \frac{d \ln L}{d\theta} = 0$$

$$\hat{\theta} = \arg \max_{\theta} L(X; \theta) \text{ ou } \hat{\theta} = \arg \max_{\theta} \ln L(X; \theta)$$

Se lit : Argument du maximum de  $L(X; \theta)$ ; (de même  $\arg \min L(X; \theta)$  se lit Argument du minimum de  $L(X; \theta)$ )

Ou  $\arg \max_{\theta} L(X; \theta)$  est la valeur de  $\theta$  pour laquelle  $L(X; \theta)$  atteint son maximum global.

**Exemple 51 :** Estimer par la méthode du Maximum de Vraisemblance l'espérance d'une loi exponentielle de paramètre  $\theta$  ;  $f(x) = \theta e^{-\theta x}$

**Etape 1 :** spécification de la fonction de *vraisemblance (log-vraisemblance)*

$$L(x_1, x_2, \dots, x_n; \theta) = f(x_1) \times f(x_1) \times f(x_1) \cdots f(x_n) = \prod_{i=1}^n f(x_i)$$

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n \theta e^{-\theta x_i} = (\theta e^{-\theta x_1}) \times (\theta e^{-\theta x_2}) \times \dots \times (\theta e^{-\theta x_n})$$

$$L(x_1, x_2, \dots, x_n; \theta) = \theta^n e^{-\theta \sum_{i=1}^n x_i} = \theta^n e^{-\theta n \bar{X}}$$

En prenant le ln de la fonction vraisemblance on a :

$$\ln L(x_1, x_2, \dots, x_n; \theta) = n \ln \theta - \theta n \bar{X}$$

**Etape 2 :** maximisation de la fonction de *vraisemblance (log-vraisemblance)*

$$\frac{d \ln L}{d\theta} = \frac{n}{\theta} - n \bar{X} = 0$$



$$T_n = \hat{\theta} = \frac{1}{\bar{X}}$$

$$\frac{d^2 \ln L}{d\theta^2} = -\frac{n}{\theta^2} < 0$$

**Exemple 52 :** Estimer par la méthode du Maximum de Vraisemblance le paramètre  $\theta$  d'une distribution de Poisson,

$$f(x) = \frac{\theta^x e^{-\theta}}{x!}$$

- **Solution :**  $\hat{\theta} = \frac{\sum X_i}{n}$

**Exemple 53 :** soit la distribution de de Pareto :

$$f(x) = \left(\frac{\alpha}{x_0}\right) \left(\frac{x}{x_0}\right)^{-(\alpha+1)} \quad \text{où } x_0 \text{ est la valeur minimum (santé)}$$

avec :

$$f(x_i) = \left(\frac{\alpha}{x_0}\right) \left(\frac{x_i}{x_0}\right)^{-(\alpha+1)}$$

Estimer par la méthode du Maximum de Vraisemblance le paramètre  $\theta$  de cette distribution

**b) la méthode log-vraisemblance concentré.**

L'exécution de la méthode de *log vraisemblance concentrée* se fait en trois (3) étapes :

**Etape 1 :** spécifiez de la fonction de *vraisemblance (log-vraisemblance)*

**Etape 2 : recherchez les extremums :** dérivée partiellement la fonction de *vraisemblance (log-vraisemblance)* par rapport à l'un des paramètres et résoudre l'équation = 0

**Etape 3 :** Insérez la solution dans l'équation de la fonction de *vraisemblance (log-vraisemblance)* de l'étape 2 et résoudre par rapport aux paramètres.

**Exemple 54 :** Estimer par la méthode de log Vraisemblance concentrée, le paramètre  $\theta$  d'une distribution de Poisson,

$$f(x) = \frac{\theta^x e^{-\theta}}{x!}$$

## 2.2. Méthode des Moments (MM)

On appelle moment d'ordre  $k$  ( $k \in \mathbb{N}^*$ ) par rapport à  $m$ , l'espérance mathématique de  $(X - m)$ :

$$\mu_k = E(X_i - m)^k$$

### 2.2.1. Moments centré et moments non centrés

Les moments non centrés d'une variable aléatoire sont les moments par rapport à  $m = 0$  :

$$m_k(X) = E(X_i)^k$$

Ainsi on aura :

- Moments centrés d'ordre  $k = 0$ :  $m_0(X) = E(X_i)^0 = 1$
- Moments centrés d'ordre  $k = 1$ :  $m_1 = E(X)$
- Moments centrés d'ordre  $k = 2$ :  $m_2 = E(X^2)$
- $\vdots$
- Moments centrés d'ordre  $k$ :  $m_k = E(X^k)$

Les moments centrés d'une variable aléatoire sont les moments par rapport à  $m = E(X)$ . On les désigne par  $\mu$  :

$$\mu_k(X) = E \left[ (X_i - E(X))^k \right]$$

- Moments centrés d'ordre  $k = 0$ :  $\mu_0 = E \left[ (X_i - E(X))^0 \right] = 1$
- Moments centrés d'ordre  $k = 1$ :  $\mu_1 = E \left[ (X_i - E(X))^1 \right] = 0$
- Moments centrés d'ordre  $k = 2$ :  $\mu_2 = E \left[ (X_i - E(X))^2 \right] = V(X)$
- ⋮
- Moments centrés d'ordre  $k$ :  $\mu_k = E \left[ (X_i - E(X))^k \right]$

NB : les moments centrés et non centrés de tous ordres n'existent pas toujours car les séries ou intégrales qui les définissent peuvent ne pas être toujours absolument convergentes. C'est pourquoi on peut parfois préciser que les moments centrés et non centrés existent jusqu'au même ordre  $k$ .

### 2.2.2. Conditions d'application de la méthode des moments

La méthode de moment requière deux conditions :

(a) **Condition de moment** : l'un des moments d'ordre  $k \in \mathbb{N}^*$ ,

- non centré :  $m_k = E_\theta(X^k) = m_k(\theta)$ , où
- centré :  $\mu_k = E_\theta \left[ (X_i - E(X))^k \right] = \mu_k(\theta)$   
doit dépendre de  $\theta$ .

(b) **Que le paramètre est estimable**

Les paramètres sont estimables par la méthode de moment si : le numéro d'ordre de moment est égale au nombre de paramètres à estimer. On dit alors que les paramètres sont justes identifiables. Dans le cas contraire, c'est-à-dire si le numéro d'ordre de moment est supérieur au nombre de paramètres à estimer (non-identifié), alors on utilise la Méthode de Moment Généralisé (GMM).

### 2.2.3. Mise en œuvre

Si les conditions d'application de la méthode des moments sont vérifiés, nous allons chercher un estimateur par résolution de l'équation en  $\theta$  obtenue en égalant moment théorique et moment empirique correspondant, soit :

$$m_{kn} = \frac{1}{n} \sum_{i=1}^n X_i^k = m_k(\theta) \text{ ou } \mu_k(\theta) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^k = \mu_k(\theta)$$

La solution de l'équation retenue, si elle existe et est unique, sera appelée estimateur obtenu par la méthode des moments.

• **Remarque :**

Dans le cas où le paramètre à estimer est  $\theta = E_\theta(X)$ , moyenne théorique de la loi, l'estimateur naturel est la moyenne empirique, ou moyenne de l'échantillon :

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



De même, pour estimer le paramètre  $\theta = V_\theta(X)$ , variance de la loi, nous retenons logiquement comme estimateur la variance empirique :

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

La variance empirique modifiée  $S_n^2$  est obtenue en divisant par  $n - 1$  au lieu de  $n$

**Exemple 55:** soit la fonction de densité

$$f(X_i) = \begin{cases} \lambda X^{\lambda-1} & \text{si } 0 \leq X \leq 1 \\ 0 & \text{si non} \end{cases}$$

- (a) Estimer  $\lambda$  par la MMV
- (b) Estimer  $\lambda$  par la MM

### 2.2.5. Propriétés

- i. La méthode des moments fournit des estimateurs convergents
- ii. La méthode des moments est conceptuellement plus simple que la méthode du maximum de vraisemblance
- iii. La méthode des moments fournit des estimateurs peu précis lorsque  $n$  est modéré
- iv. les estimateurs ainsi produits n'ont pas les bonnes propriétés asymptotiques des estimateurs du Maximum de Vraisemblance. En particulier on n'a pas en général la loi limite de l'estimateur
- v. les méthodes de : Moindre Carré Ordinaires (MCO), Maximum de Vraisemblance (MMV), Variable Instrumentale (MVI) sont des cas particuliers de la Méthode des Moments (MM).

### 2.3. Estimation ponctuelle d'une moyenne

La meilleure estimation de la moyenne  $m$  d'une population, qui puisse être déduite d'un échantillon aléatoire et simple, est la moyenne de l'échantillon.

$$\hat{m} = \bar{x} \text{ en d'autres termes } m = E(X)$$

La dispersion des différentes estimations possibles autour de cette moyenne générale, est mesurée par l'erreur standard de la moyenne :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

### 2.4. Estimation ponctuelle de la proportion (d'un pourcentage)

La meilleure estimation de la proportion  $p$  d'une population, qui puisse être déduite d'un échantillon aléatoire et simple, est la fréquence de l'échantillon  $f_n$ .

$$\hat{p} = f_n$$

La dispersion des différentes estimations possibles autour de cette proportion générale, est mesurée par l'erreur standard de la proportion :

$$\sigma_{f_n} = \sqrt{\frac{f_n(1-f_n)}{n}}$$

**Exemple 52 :** Pour un échantillon aléatoire et simple d'effectif  $n$ , dont  $x$  individus possèdent le caractère étudié, utilisé la méthode du maximum de vraisemblances pour estimer les paramètres de cette loi.

### III. ESTIMATION PAR INTERVALLE DE CONFIANCE

Afin de juger de la précision du résultat d'un problème d'estimation, nous allons associer à un échantillon d'une loi inconnue un ensemble de valeurs, et non plus une valeur unique. Ces valeurs constituent généralement un intervalle inclus dans l'ensemble  $\Theta$  des valeurs possible pour le paramètre. La détermination de cet *intervalle*, dit de *confiance*, s'effectue par la donnée de la probabilité, dénommée *niveau de confiance*, que la vraie valeur du paramètre appartienne à cet intervalle. Ainsi, l'estimation par intervalle de confiance consiste à déterminer autour de la valeur estimée un intervalle dont on a de fortes chances de croire qu'il contient la vraie valeur du paramètre recherché.

#### 3.1. Définition

Si on s'intéresse à un paramètre  $\theta$ , dont on possède un estimateur  $T$ , l'*estimation par intervalle de confiance* consiste à déterminer de part et d'autre de  $T$  les bornes  $T_1$  et  $T_2$  d'un intervalle qui a une forte probabilité de contenir  $\theta$ . Cette probabilité est appelée *niveau de confiance* et désignée par  $(1 - \alpha) \in ]0, 1[$ .  $\alpha$  est alors un risque d'erreur.

Les limites  $T_1$  et  $T_2$  sont telles que :

$$P(T_1 \leq \theta \leq T_2) = 1 - \alpha$$

L'intervalle  $[T_1, T_2]$  est appelé *intervalle de confiance (IC)*.

#### Remarque 3.1 :

La probabilité que le paramètre  $\theta$  se trouve à l'extérieur de cet intervalle est donc :

$$P(\theta \leq T_1) + P(\theta \geq T_2) = \alpha$$

#### 3.2. Principe de construction

Le risque total  $\alpha = P(\theta \leq T_1) + P(\theta \geq T_2)$  peut être réparti d'une infinité de manière.

Posons  $\alpha_1 = P(\theta \leq T_1)$  et  $\alpha_2 = P(\theta \geq T_2)$  les différents choix possibles sont les suivants :

##### 3.2.1. Intervalle unilatéral $\alpha_1 \alpha_2 = 0$

- *A droite* :  $\alpha_1 = 0, \alpha_2 = \alpha$

C'est l'interprétation donnée au paramètre  $\theta$  qui conduit à un intervalle de la forme :  $\theta > T_2$ .

- *A gauche* :  $\alpha_1 = \alpha, \alpha_2 = 0$

L'interprétation du paramètre  $\theta$  peut également conduire à un intervalle de la forme  $\theta < T_1$ .

##### 3.2.2. Intervalle bilatéral $\alpha_1 \alpha_2 > 0$

- *Symétrique*  $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ :

C'est le choix que l'on fait si la loi de  $T_n$  est symétrique, ou si on n'a aucune information particulière, ce choix étant le moins arbitraire. Dans la pratique on divise le risque  $\alpha$  en deux parties égales, Les limites  $T_1$  et  $T_2$  sont telles que :

$$P(\theta \leq T_1) = P(\theta \geq T_2) = \frac{\alpha}{2}$$

- Dissymétrique  $\alpha_1 \neq \alpha_2$ :

Seul des raisons très particulière peuvent permettre de fixer les valeurs de  $\alpha_1$  et  $\alpha_2$  telle que  $\alpha_1 + \alpha_2 = \alpha$

### 3.3. Estimation par IC d'une proportion.

Dans une population donnée de grande taille, la proportion d'individus  $p$  ayant une caractéristique donnée  $C$  est inconnue. On désire déterminer, à partir d'un tirage d'un échantillon non exhaustif de taille  $n$  de la population, un intervalle de confiance au risque  $\alpha$  de  $p$ .

Le tirage de cet échantillon peut être modélisé par un  $n$ -échantillon au hasard tiré d'une variable aléatoire  $F$  qui suit une loi de Bernoulli de paramètre  $p$ . Soient donc  $X \sim B(p)$  une loi de Bernoulli de paramètre  $p$  et  $X_1, X_2, X_3, \dots, X_n$  un  $n$ -échantillon aléatoire simple. La fréquence  $F = \frac{X_1 + X_2 + \dots + X_n}{n}$  est un bon estimateur (estimateur sans biais, convergent et efficace) du paramètre  $p$ , où chacune des variables aléatoires  $X_i$  suit une loi de Bernoulli. La fréquence est un estimateur asymptotiquement normal et on utilise l'approximation

$F \sim \mathcal{N}\left(p, \sqrt{\frac{pq}{n}}\right)$  pour  $n \geq 30, np \geq 5$  et  $nq \geq 5$ . Ces conditions seront appelées les conditions de normalités.

Les tables statistiques fournissent les valeurs  $Z_\alpha$  telles que  $P(-Z_\alpha \leq Z \leq Z_\alpha) = 1 - \alpha$  avec  $Z \sim \mathcal{N}(0, 1)$ . On applique cette relation à la variable  $Z = \frac{F-p}{\sqrt{\frac{pq}{n}}}$  qui suit une loi normale  $\mathcal{N}(0, 1)$ . On obtient

$$p \left( \left\{ -Z_\alpha < \frac{F-p}{\sqrt{\frac{pq}{n}}} < Z_\alpha \right\} \right) = 1 - \alpha$$

Remarquons que  $-Z_\alpha < \frac{F-p}{\sqrt{\frac{pq}{n}}} < Z_\alpha \Leftrightarrow -Z_\alpha < \frac{p-F}{\sqrt{\frac{pq}{n}}} < Z_\alpha \Leftrightarrow F - Z_\alpha \sqrt{\frac{pq}{n}} < p < F + Z_\alpha \sqrt{\frac{pq}{n}}$ .

on obtient un intervalle de confiance de  $p$  au niveau de confiance  $1 - \alpha$  soit  $\left[ F - Z_\alpha \sqrt{\frac{pq}{n}}; F + Z_\alpha \sqrt{\frac{pq}{n}} \right]$  pour un risque  $\alpha = 1\%$  on trouve  $Z_\alpha = 2,58$  et l'intervalle de confiance est :

$$\left[ F - 2,58 \sqrt{\frac{pq}{n}}; F + 2,58 \sqrt{\frac{pq}{n}} \right],$$

pour un risque  $\alpha = 1\%$  on trouve  $Z_\alpha = 1,96$  et l'intervalle de confiance est :

$$\left[ F - 1,96 \sqrt{\frac{pq}{n}}; F + 1,96 \sqrt{\frac{pq}{n}} \right],$$

Cet intervalle pose un problème pratique important, on peut affirmer que la proportion  $p$  appartient à cet intervalle avec une probabilité de  $1 - \alpha$  mais les bornes de cet intervalle dépendent de  $p$ , la proportion inconnue. Deux possibilités sont utilisées :

• **Première méthode :**

on remplace  $p$  et  $q$  par leurs estimations ponctuelles  $f$  et  $1 - f$ . La réalisation de l'intervalle de confiance est alors

$$\left[ f - Z_{\alpha} \sqrt{\frac{f_n(1 - f_n)}{n}}; f + Z_{\alpha} \sqrt{\frac{f_n(1 - f_n)}{n}} \right].$$

**Exemple 53 :** En vue d'un contrôle de qualité on observe la fabrication d'un objet par une machine durant une période donnée. On décide de tirer un échantillon non exhaustif de taille  $n = 1000$  dans la fabrication. On constate que 60 d'entre eux sont défectueux. Déterminer au risque de 5% un intervalle de confiance de la proportion d'objets défectueux durant la période donnée.

Les données sont  $n = 1000$  et  $f = \frac{60}{1000} = 0,06$ . La taille de l'échantillon est grande ( $n > 30$ ). Le risque de 5% conduit à  $Z_{\alpha} = 1,96$ . L'intervalle de confiance numérique est donc, au risque de 5% :

$$\left[ 0,06 - 1,96 \sqrt{\frac{0,06(1 - 0,06)}{1000}}; 0,06 + 1,96 \sqrt{\frac{0,06(1 - 0,06)}{1000}} \right] = [0,045; 0,075]$$

La proportion  $p$  d'objets défectueux fabriqués par la machine est, au risque de 5%, telle que  $4,5\% \leq p \leq 7,5\%$ .

• **Deuxième méthode :**

$pq = p(1 - p) = -p^2 + p = f(p)$ . Alors  $f'(p) = -2p + 1$  et on en déduit que  $f$  est croissante sur  $\left[0; \frac{1}{2}\right]$  et décroissante sur  $\left[\frac{1}{2}; 1\right]$ . Dans le cas où  $p$  est voisin de  $\frac{1}{2}$ , on remplace  $pq$  par sa valeur maximale  $\frac{1}{4}$ . La réalisation de l'intervalle de confiance est alors

$$\left[ f - \frac{Z_{\alpha}}{2\sqrt{n}}; f + \frac{Z_{\alpha}}{2\sqrt{n}} \right]$$

Cette méthode, qui permet un calcul rapide, donne un intervalle de confiance de grande amplitude car la valeur  $\frac{1}{4}$  du produit  $p(1 - p)$  est surestimée.

**Exemple 54 :** On a besoin d'estimer rapidement la proportion  $p$  d'accidents du travail dans une entreprise de construction. On a constaté sur un échantillon de 200 jours ouvrables qu'il y a eu 18 accidents. Déterminer, au risque de 5%, un intervalle de confiance de la proportion d'accidents.

Les données sont  $n = 200$  et  $f = \frac{18}{200} = 0,09$ . Pour un calcul rapide, l'intervalle de confiance numérique est donc, au risque de 5% :

$$\left[ 0,09 - 1,96 \frac{1}{2\sqrt{200}}; 0,09 + 1,96 \frac{1}{2\sqrt{200}} \right] = [0,02; 0,159].$$

La proportion d'accidents est au risque de 5% telle que  $2\% \leq p \leq 15,9\%$ . On se rappellera que cette méthode augmente l'amplitude de l'intervalle de confiance. Le calcul fait avec la première méthode donnerait une proportion d'accident  $p$  telle que  $5\% \leq p \leq 12,99\%$ .

**Remarque :**

- On a utilisé l'approximation normale déduite du théorème central limite pour établir l'intervalle de confiance. Il est donc nécessaire que  $n \geq 30$ ,  $np \geq 5$  et  $n(1 - p) \geq$

5. Dans la pratique,  $p$  est inconnue, on vérifie ces conditions sur  $f$  donc  $n \geq 30$ ,  $nf \geq 5$  et  $n(1 - f) \geq 5$ .

- La longueur de l'intervalle de confiance est  $L(\alpha, n) = 2Z_\alpha \sqrt{\frac{f_n(1-f_n)}{n}}$
- La précision de l'estimation obtenue est  $\frac{1}{2}L(\alpha, n) = Z_\alpha \sqrt{\frac{f_n(1-f_n)}{n}}$
- $Z_\alpha$  étant une fonction décroissante de  $\alpha$  (risque pris par le statisticien), lorsque  $1 - \alpha$  augmente,  $\alpha$  diminue,  $Z_\alpha$  augmente, la longueur de l'intervalle augmente.
- Lorsqu'on a choisi la valeur de  $\alpha$ , on peut imaginer de déterminer la taille de l'échantillon nécessaire pour atteindre une précision donnée  $l$  soit  $Z_\alpha \sqrt{\frac{f_n(1-f_n)}{n}}$ .
- On obtient  $n > Z_\alpha^2 \frac{f_n(1-f_n)}{l^2}$

**Exemple 55 :** On réalise un sondage en vue de prévoir le résultat de l'élection présidentielle. On effectue un tirage aléatoire simple de 750 électeurs. Parmi eux, 324 déclarent qu'ils ont l'intention de voter pour le candidat A tandis que 426 électeurs affirment qu'ils vont voter pour le candidat B. Donner un intervalle de confiance au niveau 95% pour la proportion d'électeurs qui vont voter pour le candidat A.

On définit la variable  $X$  de la manière suivante :

- si un électeur quelconque vote pour le candidat A,  $X = 1$  et  $p(\{X = 1\}) = p$ ,
- s'il vote pour le candidat B,  $X = 0$  et  $p(\{X = 0\}) = 1 - p$ .

$X$  suit une loi de Bernoulli de paramètre  $p$ .  $X_1, X_2, X_3, \dots, X_n$  est un échantillon aléatoire simple. Une estimation ponctuelle de  $p$  est  $f = \frac{324}{750} = 0,432$ . L'intervalle de confiance est donné par

$$\left[ f - Z_\alpha \sqrt{\frac{f_n(1-f_n)}{n}}; f + Z_\alpha \sqrt{\frac{f_n(1-f_n)}{n}} \right]$$

Les conditions de normalité sont vérifiées car  $n = 750 \geq 30$ ,  $nf = 750 \times \frac{324}{750} = 324 \geq 5$ ,  $n(1 - f) = 426 \geq 5$ . On obtient dans la table  $Z_{0,05} = 1,96$ . L'intervalle numérique est donné par :

$$\left[ 0,432 - 1,96 \sqrt{\frac{0,432 \times 0,568}{750}}; 0,432 + 1,96 \sqrt{\frac{0,432 \times 0,568}{750}} \right] = [0,39656; 0,47746]$$

On peut affirmer que  $p(\{0,39656 < p < 0,47746\}) = 0,95$  et  $p(\{p < 0,39656\}) = p(\{p > 0,47746\}) = 0,025$ .

Pour  $1 - \alpha = 0,99$  donc un risque de 1% on a  $Z_\alpha = 2,575829$ , l'intervalle de confiance est  $[0,3854; 0,4786]$ .

**Remarque 3.3.3 :** Cet exemple fait apparaître le peu d'intérêt que présente souvent les sondages tels qu'on les donne dans la presse c'est-à-dire sans donner l'intervalle de confiance ni le niveau de confiance.

2. On souhaite estimer avec une précision de 2% au niveau de confiance  $1 - \alpha = 90\%$  le pourcentage de sujets non immunisés après une vaccination. Sur combien de sujets

l'observation doit-elle porter sachant que le pourcentage observé de personnes non immunisées est

(a)  $f = 0,20$

(b)  $0.2 < f < 0,30$

(a) Supposons les conditions de normalité  $n \geq 30$ ,  $nf \geq 5$ ,  $n(1 - f) \geq 5$  vérifiées avec  $\alpha = 0,10$ ,  $f = 0,20$  et  $Z_{0,10} = 1,645$  il faut que

$$Z_{\alpha} \sqrt{\frac{f_n(1 - f_n)}{n}} \leq 0.02 \Leftrightarrow n \geq \frac{Z_{\alpha}^2 f_n(1 - f_n)}{(0,02)^2} = \frac{(1,645)^2 \times 0,2 \times 0,8}{0,0004} = 1083$$

(b) Étudions les variations de  $f(1 - f)$ . Soit  $g(x) = x(1 - x) = -x^2 + x$ ,  $g'(x) = -2x + 1$ . On en déduit alors que  $g$  est croissante sur  $[0, 2; 0, 3]$  (avec  $g(0,2) = 0,2 \times 0,8$  et  $g(0,3) = 0,3 \times 0,7$ ). Ainsi,  $0,16 < f(1 - f) < 0,21$  et par

conséquent,  $Z_{\alpha} \sqrt{\frac{f_n(1-f_n)}{n}} < Z_{\alpha} \sqrt{\frac{0,21}{n}} \leq 0.02 \Leftrightarrow n \geq \frac{(1,645)^2 \times 0,21}{0,0004} \cong 1421$ .

**3.3.1. Détermination de la taille d'échantillon n d'une IC d'une proportion.**

La question de la taille de l'échantillon joue un rôle important dans la qualité des estimations par IC lors que la population n'est pas normalement distribuée. Dans ce cas il faut choisir la taille de l'échantillon en fonction de la marge d'erreur fixée.

Si la marge d'erreur souhaitée est déterminée avant l'échantillonnage la procédure pour déterminer la taille de l'échantillon est la suivante :

Soit l'estimation par intervalle de confiance d'une proportion donnée par :

$$f \pm Z_{\alpha} \sqrt{\frac{f_n(1 - f_n)}{n}}$$

$$\bar{X} \pm Z_{\alpha} \frac{\sigma}{\sqrt{n}}$$

La quantité  $Z_{\alpha} \sqrt{\frac{f_n(1-f_n)}{n}}$   $Z_{\alpha} \frac{\sigma}{\sqrt{n}}$  correspond à la marge d'erreur. Une fois un coefficient de confiance  $1 - \alpha$  sélectionné, la valeur de  $Z_{\alpha}$  peut être déterminée. A partir des valeur de  $Z_{\alpha}$ , et  $f_n$  on peut déterminer la taille  $n$  de l'échantillon.

Soit E la marge d'erreur souhaitée

$$E = Z_{\alpha} \sqrt{\frac{f_n(1 - f_n)}{n}} = \frac{Z_{\alpha} \sqrt{f_n(1 - f_n)}}{\sqrt{n}}$$

En réarrangeant les termes de cette équation on obtient :

$$\sqrt{n} = \frac{Z_{\alpha} \sqrt{f_n(1 - f_n)}}{E}$$

$$n = \left( \frac{Z_{\alpha} \sqrt{f_n(1 - f_n)}}{E} \right)^2 = \frac{(Z_{\alpha})^2 (\sqrt{f_n(1 - f_n)})^2}{E^2}$$

$$n = \frac{(Z_{\alpha})^2 f_n(1 - f_n)}{E^2}$$



En général on fait ce calcul avant de considérer un échantillon, on n'a pas forcément de valeur pour  $f_n$ . Dans la pratique la valeur de  $f_n$  est obtenu de la manière suivante :

- Si on a estimé précédemment  $f_n$  on la considère.
- On peut utiliser une étude pilote pour déterminer  $f_n$
- L'on peut utiliser son intuition
- Si aucune de ces procédures n'est applicable, utiliser la valeur  $f^* = 0.5$

Finalement on aura :

$$n = \frac{(Z_\alpha)^2 f^* (1 - f^*)}{E^2}$$

**Exemple 56 :** Cette taille d'échantillon permet d'obtenir la marge d'erreur souhaitée au seuil de confiance choisi.

Un échantillon aléatoire simple de 800 observations généré une proportion d'échantillon  $\bar{p} = 0.70$ .

- a) Construire un intervalle de confiance à 90% pour la proportion de population
- b) Construire un intervalle de confiance à 95% pour la proportion de population

### 3.4. Intervalles associés aux paramètres de la loi normale

#### 3.4.1. Intervalles pour la moyenne d'une loi normale d'écart type connu.

La variable aléatoire  $X$  suit une loi de probabilité de paramètre  $m = E(X)$  inconnu et de variance  $\sigma^2$  connue. Soit  $X_1, X_2, X_3, \dots, X_n$  un  $n$ -échantillon aléatoire simple  $X$ . on sait alors qu'un bon estimateur ponctuel de  $m$  est  $\bar{X}$  (estimateur sans biais, convergent et efficace) et que

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \sim \mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right) \text{ et } Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$$

Les tables fournissent les valeurs  $Z_\alpha$  pour  $\alpha$  donné, telle que telles que

$$p(-Z_\alpha \leq Z \leq Z_\alpha) = 1 - \alpha$$

or

$$\begin{aligned} -Z_\alpha \leq Z \leq Z_\alpha &\Leftrightarrow -Z_\alpha < \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} < Z_\alpha \\ \Leftrightarrow -Z_\alpha \frac{\sigma}{\sqrt{n}} < \bar{X} - m < Z_\alpha \frac{\sigma}{\sqrt{n}} &\Leftrightarrow -Z_\alpha \frac{\sigma}{\sqrt{n}} < m - \bar{X} < Z_\alpha \frac{\sigma}{\sqrt{n}} \end{aligned}$$

On obtient ainsi

$$p\left(\left\{\bar{X} - Z_\alpha \frac{\sigma}{\sqrt{n}} < m < \bar{X} + Z_\alpha \frac{\sigma}{\sqrt{n}}\right\}\right) = 1 - \alpha$$

C'est-à-dire un intervalle de confiance de  $m$  au niveau de confiance  $1 - \alpha$  soit

$$\left[\bar{X} - Z_\alpha \frac{\sigma}{\sqrt{n}}; \bar{X} + Z_\alpha \frac{\sigma}{\sqrt{n}}\right].$$

Dans la pratique, on dispose d'un échantillon non exhaustif tiré au hasard de la population. Cet échantillon fournit une réalisation de  $X$  par le calcul de la moyenne  $\bar{x}$ . Ainsi l'échantillon donne une réalisation de l'intervalle de confiance au risque  $\alpha$  qui est

$$\left[\bar{x} - Z_\alpha \frac{\sigma}{\sqrt{n}}; \bar{x} + Z_\alpha \frac{\sigma}{\sqrt{n}}\right]$$

**Exemple 57 :** Une machine  $M$  fabrique des engrenages en grande série. Des études antérieures permettent de dire que les mesures des diamètres forment une population normale d'écart-type  $\sigma = 0,042$

cm. On extrait un échantillon non exhaustif de la fabrication journalière de taille  $n = 200$  engrenages.

La moyenne des diamètres sur cet échantillon est  $\bar{x} = 0,824$  cm. Donner au seuil de confiance 95% un intervalle de confiance de la moyenne  $m$  des diamètres des engrenages.

Considérons  $D$  la variable aléatoire égale au diamètre des engrenages. L'énoncé dit que  $D \sim \mathcal{N}(m, \sigma = 0,042)$ . Soit  $D_1, D_2, \dots, D_{200}$  un 200-échantillon au hasard de  $D$ . Les  $n = 200$  variables aléatoires  $D$  suivent la même loi  $\mathcal{N}(m, \sigma = 0,042)$  que  $D$ . Soit  $m$  le diamètre moyen inconnu des engrenages. On considère alors l'estimateur sans biais et convergent  $\bar{D} = \frac{1}{200} \sum_{i=1}^{200} D_i$  de  $m$ . une réalisation de  $\bar{D}$  est  $d = 0,824$ . On sait que l'intervalle de confiance au risque  $\alpha$  est  $\left[ \bar{D} - Z_\alpha \frac{\sigma}{\sqrt{n}}; \bar{D} + Z_\alpha \frac{\sigma}{\sqrt{n}} \right]$ . Pour un risque de 5% on a  $Z_\alpha = 1,96$ . Ainsi l'intervalle de confiance est  $\left[ \bar{D} - 1,96 \frac{0,042}{\sqrt{200}}; \bar{D} + 1,96 \frac{0,042}{\sqrt{200}} \right]$ . L'échantillon fournit une réalisation de cet intervalle de confiance à savoir  $\left[ 0,824 - 1,96 \frac{0,042}{\sqrt{200}}; 0,824 + 1,96 \frac{0,042}{\sqrt{200}} \right]$ . Soit savoir  $[0,818; 0,830]$

**Exemple 58 :** La moyenne d'un échantillon aléatoire simple de 40 éléments est égale à 25. L'écart-type de la population est  $\sigma = 5$ .

- a) Quelle est l'erreur type de la moyenne.
- b) Pour un seuil de 95%, quelle est la marge d'erreur

**3.4.2. Intervalles pour la moyenne d'une loi normale d'écart type inconnu.**

Dans la pratique, si l'espérance  $m = E(X)$  est inconnue, a fortiori, la variance  $\sigma^2 = E[(X - m)^2]$  est également inconnue. Or nous venons de voir que l'intervalle de confiance de  $m$  tel qu'il vient d'être défini dépend de  $\sigma$ . Il est alors tentant de remplacer  $\sigma$  par son estimation ponctuelle  $s$  fournie par l'estimateur  $S^2$ . Ce nombre n'est autre que l'écart-type calculé sur l'échantillon de taille  $n$  avec  $(n - 1)$  degrés de liberté (ddl). Dans ces conditions, on utilise le procédé dit de « *Studentisation* » qui consiste à remplacer la variable centrée réduite  $Z = \frac{X - E(\bar{X})}{\sigma(\bar{X})}$  par la variable  $T = \frac{X - E(\bar{X})}{\frac{S}{\sqrt{n}}}$  qui suit une loi de Student à  $n - 1$  ddl. La table de Student

nous permet de déterminer  $t_{n-1, \alpha}$  tel que pour  $(n - 1)$  ddl on ait

$$p(-t_{n-1, \alpha} < T < t_{n-1, \alpha}) = 1 - \alpha$$

On obtiendra alors l'intervalle de confiance au risque  $\alpha$  :

$$\left[ \bar{X} - t_{(n-1, \alpha)} \frac{S}{\sqrt{n}}; \bar{X} + t_{(n-1, \alpha)} \frac{S}{\sqrt{n}} \right]$$

Dont une réalisation sur l'échantillon sur l'échantillon est  $\left[ \bar{x} - t_{(n-1, \alpha)} \frac{s}{\sqrt{n}}; \bar{x} + t_{(n-1, \alpha)} \frac{s}{\sqrt{n}} \right]$

**Exemple 59 :** Dans l'atmosphère, le taux de gaz nocif, pour un volume donné, suit une loi normale d'espérance et de variance inconnues. On effectue  $n$  prélèvements conduisant aux valeurs numériques  $x_1, x_2, x_3, \dots, x_n$

- (a) Sur un échantillon de taille  $n = 10$ , on observe  $\bar{x} = 50$  et  $s^2 = 100$ .  
 Quel est l'intervalle de confiance à 5% du taux moyen  $m$  de gaz dans l'atmosphère ?
- (b) Quel serait cet intervalle si la variance  $\sigma^2$  du taux de gaz nocif était connue et valait exactement 100 ?

(a) Considérons  $X$  la variable aléatoire égale au taux de gaz nocif dans l'atmosphère. L'énoncé dit que  $X \sim N(m, \sigma)$  avec  $m$  et  $\sigma^2$  inconnues ( $m$  représente le taux moyen de gaz nocif dans l'atmosphère). Soit  $X_1, X_2, X_3, \dots, X_n$  un  $n$ -échantillon au hasard de  $X$ . Les  $n$  variables aléatoires  $X_i$  suivent la même loi  $N(m, s)$  que  $X$ . Les valeurs observées  $x_1, x_2, x_3, \dots, x_n$  sont une réalisation du  $n$ -échantillon de  $X$ . On considère alors l'estimateur sans biais et convergent  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  de  $m$ .

Une réalisation de  $\bar{X}$  sur l'échantillon est  $\bar{x} = 50$ . L'intervalle de confiance de  $m$  au risque  $\alpha$  est l'intervalle aléatoire  $\left[ \bar{X} - Z_\alpha \frac{\sigma}{\sqrt{n}}; \bar{X} + Z_\alpha \frac{\sigma}{\sqrt{n}} \right]$   $Z_\alpha$  est déterminé par  $p(-Z_\alpha < T < Z_\alpha) = 1 - \alpha$  avec  $T \sim N(0, 1)$ .

Or  $\sigma$  est inconnu, on le remplace donc par son estimation  $s$ . L'intervalle de confiance sur l'échantillon devient :

$$\left[ \bar{X} - t_{(n-1, \alpha)} \frac{S}{\sqrt{n}}; \bar{X} + t_{(n-1, \alpha)} \frac{S}{\sqrt{n}} \right]$$

Où  $t_{(n-1, \alpha)}$  est déterminé  $p(-t_{(n-1, \alpha)} < T < t_{(n-1, \alpha)}) = 1 - \alpha$  avec  $T$  qui suit une loi de Student à  $n - 1$  ddl.

Pour  $\alpha = 5\%$  et  $n - 1 = 9$ , on obtient dans la table  $t_{(9, 0,05)} = 2,262$ . L'intervalle de confiance recherché est donc :

$\left[ \bar{X} - 2,262 \frac{S}{\sqrt{n}}; \bar{X} + 2,262 \frac{S}{\sqrt{n}} \right]$ . Une réalisation de cet intervalle de confiance sur l'échantillon est :  $\left[ \bar{X} - 2,262 \frac{s}{\sqrt{n}}; \bar{X} + 2,262 \frac{s}{\sqrt{n}} \right]$  soit  $\left[ 50 - 2,262 \frac{10}{\sqrt{10}}; 50 + 2,262 \frac{10}{\sqrt{10}} \right] = [42,84; 57,15]$  numériquement, qui est donc l'intervalle de confiance, au risque 5% du taux moyen du gaz nocif dans l'atmosphère.

(b) Si la variance est connue et égale à 100, on utilise la table de la loi normale pour déterminer  $Z_{0,05} = 1.96$ . L'intervalle de confiance est alors  $\left[ \bar{X} - 1.96 \frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}} \right]$  et une réalisation sur l'échantillon est :  $\left[ 50 - 1.96 \frac{10}{\sqrt{10}}; 50 + 1.96 \frac{10}{\sqrt{10}} \right] = [43,80; 56,20]$ .

L'utilisation de la loi normale donne un intervalle de confiance d'amplitude plus petit que celui obtenu à l'aide de la loi de Student. Ce résultat est cohérent car pour l'utilisation de la loi normale, on a supposé que l'écart-type  $\sigma$  était connu. On a donc une meilleure connaissance de la loi de  $X$  que dans le cas où  $\sigma$  est inconnu.

**Remarque 3.3.4** Si la taille de l'échantillon est "grande" ( $n > 30$ ), on peut utiliser la loi normale à la place de la loi de Student. C'est pour cette raison qu'on trouve dans la littérature l'expression : "la loi de Student est la loi des petits échantillons".

### 3.4.3. Détermination de la taille de l'échantillon pour une population qui ne suit une pas une distribution normale.

Considérons la question de la taille de l'échantillon nécessaire pour estimer avec un niveau de précision donné la proportion de la population. Le raisonnement suivi pour déterminer la taille de l'échantillon impliqué dans la construction d'un IC pour  $\bar{X}$  est similaire à celui suivi dans le cas de la proportion.

Soit  $E$  la marge d'erreur souhaitée

$$E = Z_{\alpha} \frac{\sigma}{\sqrt{n}}$$

En résolvant cette équation on obtient :

$$\begin{aligned}\sqrt{n} &= \frac{Z_{\alpha}\sigma}{E} \\ n &= \left(\frac{Z_{\alpha}\sigma}{E}\right)^2 = \frac{(Z_{\alpha})^2\sigma^2}{E^2} \\ \mathbf{n} &= \frac{(Z_{\alpha})^2\sigma^2}{E^2}\end{aligned}$$

Cette taille d'échantillon permet d'obtenir la marge d'erreur souhaitée au seuil de confiance choisi.

**Exemple 60 :** Quelle est doit être la taille de l'échantillon pour obtenir un intervalle de confiance de 95% avec une marge d'erreur de 10 ? Supposer que l'écart type de la population est égale à 40.

#### 3.4.4. Intervalles pour la variance d'une loi normale d'espérance connue.

L'estimateur sans biais de  $\sigma^2$ , basé sur l'échantillon  $(X_1, \dots, X_n)$  de cette loi  $N(m, \sigma)$  où  $m$  est connu est :

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$$

De loi connue :  $\frac{n\hat{\sigma}_n^2}{\sigma^2} \sim \chi_n^2$ . On peut donc déterminer les valeurs de  $t_1$  et  $t_2$  telles que :

$$p\left(t_1 < n \frac{\hat{\sigma}_n^2}{\sigma^2} < t_2\right) = 1 - \alpha$$

Ce qui conduit à l'intervalle de confiance défini par :

$$p\left(n \frac{\hat{\sigma}_n^2}{t_2} < \sigma^2 < n \frac{\hat{\sigma}_n^2}{t_1}\right) = 1 - \alpha$$

Cependant, il n'y a qu'une seule condition pour déterminer les deux valeurs  $t_1$  et  $t_2$  et il reste un degré d'incertitude puisque la loi utilisée n'est pas symétrique. Si on pose  $\alpha_1 =$

$P(\chi_n^2 < t_1)$  et  $\alpha_2 = P(\chi_n^2 < t_2)$ , la seule contrainte dans le choix de  $\alpha_1$  et  $\alpha_2$  est  $\alpha_1 + \alpha_2 = \alpha$

#### 3.4.5. Intervalles pour la variance d'une loi normale d'espérance inconnue.

Quand le second paramètre  $m$  de la loi normale est inconnu, l'estimateur sans biais de  $\sigma^2$  qu'il faut retenir est :

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

Sa loi est connue,  $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2$ . On peut donc déterminer les valeurs de  $t_1$  et  $t_2$  telles que :

$$p\left(t_1 < \frac{(n-1)S_n^2}{\sigma^2} < t_2\right) = 1 - \alpha$$

Ce qui conduit à l'intervalle de confiance défini par :

$$p\left((n-1) \frac{S_n^2}{t_1} < \sigma^2 < (n-1) \frac{S_n^2}{t_2}\right) = 1 - \alpha$$

Cependant, il n'y a qu'une seule condition pour déterminer les deux valeurs  $t_1$  et  $t_2$  et il reste un degré d'incertitude puisque la loi utilisée n'est pas symétrique.

Si on pose  $\alpha_1 = P(\chi_{n-1}^2 < t_1)$  et  $\alpha_2 = P(\chi_{n-2}^2 < t_2)$ , la seule contrainte est  $\alpha_1 + \alpha_2 = \alpha$

#### IV. CONCLUSION

Dans ce chapitre nous avons donc montré comment utiliser un échantillon pour développer des estimations ponctuelles et par intervalle de confiance des paramètres d'une population. La comparaison des paramètres issus de deux échantillons n'ont pas été traitée dans ce chapitre, toutefois pour compléter ses connaissances l'étudiant pourra se rapporter aux abondantes documentations qui existent dans le domaine et tout particulièrement au livre intitulé « *statistique pour l'économie et la gestion* » d'Anderson, Sweeney et Williams (chapitre 10 et 11). Dans le chapitre suivant nous poursuivrons notre découverte de la statistique inférentielle en étudiant les tests d'hypothèse afin de déterminer si une assertion au sujet de la valeur d'un paramètre de la population doit être ou non rejetée.

# C HAPITRE VI

## TEST D'HYPOTHESES

Un test d'hypothèse est une méthode permettant de prendre une décision (accepter ou rejeter) à partir de l'étude d'un ou plusieurs échantillons aléatoires, Les tests statistiques ou les tests d'hypothèses ont donc pour but de vérifier, à partir de données observées dans un ou plusieurs échantillons, la validité de certaines hypothèses relatives à une ou plusieurs populations. Les méthodes de l'inférence statistique nous permettent de déterminer, avec une probabilité donnée, si les différences constatées au niveau des échantillons peuvent être imputables au hasard ou si elles sont suffisamment importantes pour signifier que les échantillons proviennent de populations vraisemblablement différentes.

On distinguera deux classes de tests :

- *Les tests paramétriques*, requièrent un modèle à forte contraintes (normalité des distributions ou approximation normale pour des grands échantillons). Quand ses conditions sont remplies, ils sont plus puissants que les tests non paramétriques.
- *Les tests non paramétriques* s'emploient lorsque les conditions d'applications des autres méthodes ne sont pas satisfaites, même après d'éventuelles transformations de variables. Ils peuvent s'utiliser même pour des échantillons de taille très faible. Il n'y a pas d'hypothèse de normalité au préalable.

*Les tests paramétriques*, quand ses conditions sont remplies, sont plus puissants que les *tests non paramétriques*.

### I. GENERALITES SUR LES TESTS STATISTIQUES

#### 1.1.Principe d'un test statistique

Pour effectuer un test d'hypothèses, on commence par faire une hypothèse sur un paramètre de la population considérée. Cette hypothèse est appelée *hypothèse nulle* et est notée  $H_0$ . On définit ensuite une autre hypothèse, appelée *hypothèse alternative* notée  $H_1$  ou  $H_a$  qui correspond à l'opposé de ce qui est établi dans l'*hypothèse nulle*.

La procédure de test consiste alors à utiliser les données issues d'un échantillon pour tester les deux assertions en compétition et l'on conclut en acceptant ou en rejetant l'hypothèse de travail à partir de règles de décisions objectives.

##### 1.1.1. Méthodologie.

Différentes étapes doivent être suivies pour tester une hypothèse :

- 1) définir les hypothèses nulle  $H_0$ , et alternative  $H_a$ ;
- 2) choisir une statistique pour contrôler  $H_0$ ;
- 3) définir le seuil de signification du test  $\alpha$  et la région critique associée ;

- 4) calculer, à partir des données fournies par l'échantillon, la valeur de la statistique ;
- 5) prendre une décision concernant l'hypothèse posée.

### 1.1.2. Définir les hypothèses nulle et alternative.

Définir les hypothèses de travail, constitue un élément essentiel des tests d'hypothèses de même que vérifier les conditions d'application de ces dernières. Dans certain cas, il n'est pas évident de formuler *les hypothèses nulle* et *alternative*. C'est pourquoi il est important d'être très attentif à la formulation des hypothèses afin d'être sûr qu'elles sont appropriées et que les conclusions du test d'hypothèse fournissent bien les informations souhaitées par le chercheur ou le responsable.

*L'hypothèse nulle* notée  $H_0$  est l'hypothèse que l'on désire contrôler : elle consiste à dire qu'il n'existe pas de différence entre les paramètres comparés ou que la différence observée n'est pas significative et est due aux fluctuations d'échantillonnage. Cette hypothèse est formulée dans le but d'être rejetée.

*L'hypothèse alternative* notée  $H_1$  est la "négation" de  $H_0$ , elle est équivalente à dire « $H_0$  est fausse». La décision de rejeter  $H_0$  signifie que  $H_1$  est réalisée ou  $H_1$  est vrai.

**Remarque :** Il existe une dissymétrie importante dans les conclusions des tests. En effet, *la décision d'accepter  $H_0$  n'est pas équivalente à «  $H_0$  est vraie et  $H_1$  est fausse »*. Cela traduit seulement l'opinion selon laquelle, *il n'y a pas d'évidence nette pour que  $H_0$  soit fausse*.

**NB :** Un test conduit à rejeter ou à ne pas rejeter une hypothèse nulle jamais à l'accepter d'emblée.

La nature de  $H_0$  détermine la façon de formuler  $H_1$  et par conséquent la nature unilatérale ou bilatérale du test.

Dans certains cas, il n'est pas évident de formuler les hypothèses nulle et alternative. Il faut donc être très attentif à la formulation des hypothèses, afin d'être sûr qu'elles sont appropriées et que les conclusions du test d'hypothèses fournissent bien les informations souhaitées par le chercheur ou le responsable.

On se placera dans les cas suivants :

#### a) Test unilatéral

Dans des situations nécessitant de tester la validité d'une assertion, l'hypothèse nulle correspond généralement à l'hypothèse selon laquelle l'assertion est vraie. L'hypothèse alternative est ensuite formulée de manière à ce que le rejet de  $H_0$  prouve statistiquement que l'hypothèse établie est correcte. On peut alors corriger l'assertion lorsque  $H_0$  est rejeté. Ainsi on parle de *test unilatéral* lorsque l'hypothèse alternative se "*compose d'une seule partie*".  $H_1$  et peut prendre l'une des valeurs suivantes :

- *Test unilatéral à gauche*

Très utiliser pour tester la validité d'une assertion, dans ce test L'hypothèse nulle correspond généralement à l'hypothèse selon laquelle l'assertion est vraie et présente comme suit :

$$H_0: \theta \geq \theta_0$$

$$H_1: \theta < \theta_0$$

**Exemple 61 :** On considère un producteur de boisson non alcoolisée qui prétend que les bouteilles de deux litres contiennent en moyenne 2,028 litres de son produit, au minimum. Un échantillon de bouteilles de deux litres est sélectionné et leur capacité de contenance est

mesurée pour tester l'affirmation du fabricant. Les hypothèses nulle et alternative sont formulées de la façon suivante :

$$H_0 : \mu \geq 2,028$$

$$H_1 : \mu > 2,028$$

avec  $\mu$  désignant la capacité de contenance moyenne mesurée.

- **Test unilatéral à droite**

Souvent utilisé pour tester les hypothèses de recherche, les hypothèses nulle et alternative doivent être formulées de manière à ce que le rejet de  $H_0$  conduise à la conclusion souhaitée. Elle se présente de la manière suivante :

$$H_0 : \theta \leq \theta_0$$

$$H_1 : \theta > \theta_0$$

**Exemple 62 :** On considère un modèle de voiture particulier qui consomme, en moyenne, un litre d'essence tous les 24 kilomètres. Un groupe de recherche a mis au point un nouveau moteur spécialement conçu pour augmenter le nombre de kilomètres effectués avec un litre d'essence. Pour évaluer les performances du nouveau moteur, plusieurs exemplaires du prototype ont été construits, installés sur des voitures et soumis à des tests de conduite par les groupes de recherche (le groupe de recherche cherche donc à prouver que le nouveau moteur augmente en moyenne le nombre de kilomètres effectués avec un litre d'essence). Par conséquent, les hypothèses nulles et alternative appropriées à cette étude sont :

$$H_0 : \mu \leq 24$$

$$H_1 : \mu > 24$$

Avec  $\mu$  désignant la distance moyenne parcourue avec un litre d'essence par le nouveau moteur.

**Exemple 63 :** si l'on fait l'hypothèse que la fréquence de fumeurs dans la population estudiantine  $p$  est supérieure à la fréquence de fumeurs dans la population  $p_0$ , on pose alors

$H_0 : p = p_0$  et  $H_1 : p > p_0$ . Le test sera unilatéral à droite car on considère que la fréquence  $p$  ne peut être que supérieure à la fréquence  $p_0$ .

Il aurait été possible également d'avoir :  $H_0 : p = p_0$  et  $H_1 : p < p_0$  (cas de test unilatéral à gauche)

### b) Test bilatéral : Tester des hypothèses dans un contexte de prise de décision

Dans les deux cas précédent des mesures sont prises lorsque  $H_0$  est rejetée. Cependant dans de nombreux cas, des mesures doivent être prises à la fois quand  $H_0$  n'est pas rejetée et quand  $H_0$  est rejetée. En générale ce type de situation se présente **dans un contexte de prise de décision** où l'on doit choisir entre deux actions l'une associée à l'hypothèse nulle et l'autre associée à l'hypothèse alternative. Le **test bilatéral** se "décompose en deux parties" :

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

**Exemple 64 :** Sur la base d'un échantillon de pièces, issues d'une livraison qui vient d'être reçue, un inspecteur de contrôle de la qualité doit décider d'accepter ou de refuser la livraison, en fonction des critères de qualité établis. Supposons que les critères de qualité d'une pièce particulière correspondent à une longueur moyenne de deux centimètres. Si la moyenne est supérieure ou inférieure à deux centimètres, les pièces poseront problème dans le processus d'assemblage. Dans ce cas les hypothèses nulle et alternative sont formulées de la façon suivante :

$$H_0: \mu = 2$$

$$H_1: \mu \neq 2$$

avec  $\mu$  désignant la mesure moyenne en centimètres des pièces de la livraison.

**Exemple 65** : si  $H_0$  consiste à dire que la population estudiantine avec une fréquence de fumeurs  $p$  est représentative de la population globale avec une fréquence de fumeurs  $p_0$ , on pose alors :

$$H_0: p = p_0 \quad H_1: p \neq p_0$$

Le test sera bilatéral car on considère que la fréquence  $p$  peut être supérieure ou inférieure à la fréquence  $p_0$ .

### 1.2. Choix d'un test statistique

Ce choix dépend de la nature des données, du type d'hypothèse que l'on désire contrôler, des affirmations que l'on peut admettre concernant la nature des populations étudiées (normalité, égalité des variances) et d'autres critères que nous préciserons.

Un *test statistique* ou une statistique est une fonction des variables aléatoires représentant l'échantillon dont la valeur numérique obtenue pour l'échantillon considéré permet de distinguer entre  $H_0$  vraie et  $H_0$  fausse.

Dans la mesure où la loi de probabilité suivie par le paramètre  $p_0$  au niveau de la population en général est connue, on peut ainsi établir la loi de probabilité de la statistique  $S$  telle que :

$$S = p - p_0$$

(Voir intervalle de confiance d'une fréquence)

### 1.3. Choix de la région critique

Une statistique est une fonction des variables aléatoires représentant l'échantillon. Le choix de la statistique dépend de la nature des données, du type d'hypothèse que l'on désire contrôler, des affirmations que l'on peut admettre concernant la nature des populations étudiées .... La valeur numérique de la statistique obtenue pour l'échantillon considéré permet de distinguer entre  $H_0$  vraie et  $H_0$  fausse.

Connaissant la loi de probabilité suivie par la statistique  $S$  sous l'hypothèse  $H_0$ , il est possible d'établir une valeur seuil,  $S_{seuil}$  de la statistique pour une probabilité donnée appelée le niveau de signification  $\alpha$  du test.

La région critique  $R_c = f(S_{seuil})$  correspond à l'ensemble des valeurs telles que :

$$\mathbb{P}(S \in R_c) = \alpha \Rightarrow S > S_{seuil}$$

Et le niveau de signification est telle que :

$$P(S > S_{seuil}) = \alpha \text{ avec } P(S \leq S_{seuil}) = 1 - \alpha$$

Selon la nature unilatérale ou bilatérale du test, la définition de la région critique varie.

Hypothèse alternative	Test unilatéral $H_0: t = t_0$		Test Bilatéral $H_0: t = t_0$
Hypothèse alternative	$H_1: t > t_0:$	$H_1: t < t_0$	$H_1: t \neq t_0$
Valeur de S sous H1 $S = t - t_0$	$S > 0$	$S < 0$	$ S  \neq 0$
Niveau de signification	$P(S > S_{seuil}) = \alpha$	$P(S < S_{seuil}) = \alpha$	$P( S  > S_{seuil}) = \alpha$

**1.4. Règle de décision**

Il existe deux stratégies pour prendre une décision en ce qui concerne un test d'hypothèse :

- ✓ la première stratégie fixe **à priori** la valeur du seuil de signification  $\alpha$
- ✓ la seconde établit la valeur de la probabilité critique  $\alpha_{obs}$  **à posteriori**.

- **Règle de décision 1 :**

Sous l'hypothèse «  $H_0$  est vraie » et pour *un seuil de signification  $\alpha$  fixé*

- Si la valeur de la statistique S calculée ( $S_{obs}$ ) est supérieure à la valeur seuil  $S_{seuil}$  (appartient à la région critique) :

$$S_{obs} > S_{seuil}$$

alors l'hypothèse  $H_0$  est rejetée au risque d'erreur  $\alpha$  et l'hypothèse  $H_1$  est acceptée ;

- Si la valeur de la statistique S calculée ( $S_{obs}$ ) est inférieure à la valeur seuil  $S_{seuil}$  (n'appartient pas à la région critique)

$$S_{obs} < S_{seuil}$$

alors l'hypothèse  $H_0$  ne peut être rejetée.

**Remarque :** Le choix du niveau de signification ou risque  $\alpha$  est lié aux conséquences pratiques de la décision, on choisira  $\alpha = 1\%$  ou  $1\%$ , mais si le débat est plutôt académique, le traditionnel  $\alpha = 5\%$  fera le plus souvent l'affaire.

- **Règle de décision 2 :**

La probabilité critique  $\alpha$  telle que  $P(S \geq S_{obs}) = \alpha_{obs}$  est évaluée

- Si  $\alpha_{obs} \geq \alpha = 0,05$  l'hypothèse  $H_0$  est acceptée car le risque d'erreur de rejeter  $H_0$  alors qu'elle est vrai est trop important ;
- Si  $\alpha_{obs} < \alpha = 0,05$  l'hypothèse  $H_0$  est rejetée car le risque d'erreur de rejeter  $H_0$  alors qu'elle est vrai est très faible.

**1.5. Risque d'erreur, puissance et robustesse d'un test**

Les hypothèses alternative et nulle sont les assertions opposées au sujet d'un paramètre de la population. Soit l'hypothèse nulle  $H_0$  est vraie, soit l'hypothèse alternative  $H_a$  est vraie, mais pas les deux. Idéalement la procédure de test devrait conduire à l'acceptation de  $H_0$  lorsque  $H_0$  est vraie et au rejet de  $H_0$  lorsque  $H_a$  est vraie. Malheureusement, ce résultat idéal n'est pas toujours obtenu. Puisque les tests d'hypothèses sont basés sur les informations d'un échantillon, nous devons admettre la possibilité d'erreurs. Le tableau illustre les deux types d'erreurs qui peuvent survenir dans un test d'hypothèse.



décision \ réalité	$H_0$ est vraie	$H_0$ est fausse
accepter $H_0$	$1 - \alpha$	erreur de type II ( $\beta$ )
rejeter $H_0$	erreur de type I ( $\alpha$ )	$1 - \beta$ (puissance)

Le risque d'erreur  $\alpha$  est la probabilité que la valeur expérimentale ou calculée de la statistique  $S$  appartienne à la région critique si  $H_0$  est vrai. Dans ce cas  $H_0$  est rejetée et  $H_1$  est considérée comme vraie.

### 1.5.1. Risque d'erreur de première espèce

Le *seuil critique* ou *seuil de signification* du test, notée  $\alpha$  est la probabilité de faire une erreur de première espèce lorsque l'hypothèse nulle est vraie et satisfaite avec égalité.

On appelle *erreur de première espèce* la probabilité de rejeter  $H_0$  et d'accepter  $H_1$  alors que  $H_0$  est vraie.

$$\alpha = P(\text{rejeter } H_0 / H_0 \text{ est vraie})$$

Ou accepter  $H_1$  alors qu'elle est fausse.

$$\alpha = P(\text{accepter } H_1 / H_1 \text{ fausse})$$

Ceci se produit si la valeur de la statistique de test tombe dans la région de rejet alors que l'hypothèse  $H_0$  est vraie. La probabilité de cet événement est le niveau de signification  $\alpha$ . On dit aussi que le niveau de signification est la probabilité de rejeter l'hypothèse nulle à tort.

**Remarque :** La valeur du risque  $\alpha$  doit être fixée a priori par l'expérimentateur et jamais en fonction des données. C'est un compromis entre le risque de conclure à tort et la faculté de conclure.

Toutes choses étant égales par ailleurs, la région critique diminue lorsque  $\alpha$  décroît (voir intervalle de confiance) et donc on rejette moins fréquemment  $H_0$ . A vouloir commettre moins d'erreurs, on conclut plus rarement.

**Exemple 63:** Si l'on cherche à tester l'hypothèse qu'une pièce de monnaie n'est pas « truquée », nous allons adopter la règle de décision suivante :

$H_0$  : la pièce n'est pas truquée

- est acceptée si  $X \in [40; 60]$
- rejetée si  $X \notin [40; 60]$  donc soit  $X < 40$  ou  $X > 60$  avec  $X$  « nombre de face » obtenus en lançant 100 fois la pièce. Le risque d'erreur de première espèce est  $\alpha = P(B(100; 1/2) \in [40; 60])$ .

### 1.5.2. Risque d'erreur de deuxième espèce

Le **risque d'erreur**  $\beta$  est la probabilité que la valeur expérimentale ou calculée de la statistique n'appartienne pas à la région critique si  $H_1$  est vrai. Dans ce cas  $H_0$  est acceptée et  $H_1$  est considérée comme fausse.

On appelle *erreur de seconde espèce*, notée  $\beta$  la probabilité de rejeter  $H_1$  et d'accepter  $H_0$  alors que  $H_1$  est vraie.

$$\beta = P(\text{accepter } H_0 / H_0 \text{ est fausse}) \text{ ou } P(\text{accepter } H_0 / H_1 \text{ est vraie})$$

Ou rejeter  $H_1$  alors qu'elle est vraie

$$\beta = P(\text{rejeter } H_1 / H_1 \text{ est vraie})$$

**Remarque :** Pour quantifier le risque  $\beta$ , il faut connaître la loi de probabilité de la statistique

S sous l'hypothèse  $H_1$ .

**Exemple 64:** Si l'on reprend l'exemple précédent de la **pièce de monnaie**, la probabilité  $p$  d'obtenir face est de 0,6 pour une pièce truquée. Si l'on adopte toujours la même règle de décision :

$H_0$  : la pièce n'est pas truquée est :

- ✓ acceptée si  $X \in [40; 60]$
- ✓ rejetée si  $X \notin [40; 60]$  donc soit  $X < 40$  ou  $X > 60$  avec  $X$  « nombre de face » obtenus en lançant 100 fois la pièce.

Quel est le risque d'erreur de seconde espèce  $\beta$  dans ce cas ?

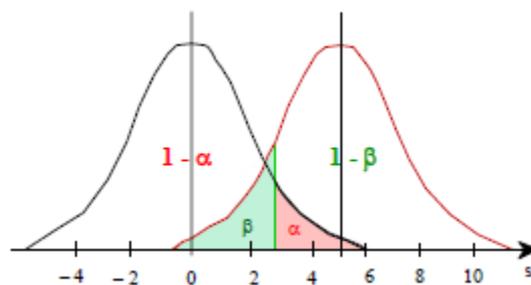
**1.5.3. Puissance et robustesse d'un test ( $1 - \beta$ )**

Les tests ne sont pas faits pour « démontrer »  $H_0$  mais pour « rejeter »  $H_0$ . L'aptitude d'un test à rejeter  $H_0$  alors qu'elle est fautive constitue **la puissance du test**.

La **puissance d'un test** est donc la probabilité de rejeter  $H_0$  alors qu'elle est fautive, ou de d'accepter  $H_1$  alors qu'elle est vraie:

$$1 - \beta = P(\text{rejeter } H_0 / H_0 \text{ fautive}) = P(\text{accepter } H_1 / H_1 \text{ vraie})$$

La relation entre les deux risques d'erreur figure sur le graphe ci-dessous.



**Remarque :**

- ✓ La puissance d'un test est fonction de la nature de  $H_1$ , un test unilatéral est plus puissant qu'un test bilatéral.
- ✓ La puissance d'un test augmente avec la taille de l'échantillon  $N$  étudié à valeur de  $\alpha$  constant.
- ✓ La puissance d'un test diminue lorsque  $\alpha$  diminue.

**Exemple 65 :** Si l'on reprend l'exemple précédent de la pièce de monnaie, calculez la puissance du test lorsque la probabilité d'obtenir face est respectivement 0,3 - 0,4 - 0,6 - 0,7 - 0,8 pour une pièce truquée. Que constatez-vous ?

Les différentes situations que l'on peut rencontrer dans le cadre des tests d'hypothèse sont résumées dans le tableau suivant :

Décision \ Réalité	Ho vraie	Ho fautive
	Non-rejet de Ho	Correct
Rejet de Ho	Rejet à tort Risque de première espèce $\alpha$	Puissance du test $1 - \beta$



## BIBLIOGRAPHIE SELECTIVE

**Anderson D.R., Sweeney D. J., Williams T. A., 2007 :** *Statistique pour l'économie et la gestion* » 2<sup>ème</sup> éd. De Boeck, Bruxelles. 803p

**Labrousse C., 1978 :** *Statistique : exercices corrigés avec rappel de cours. Tome 2, sciences économiques* 2<sup>ème</sup> année. 4<sup>ème</sup> édition. Dunod, Paris. 282p

**Lecoutre J., P., 2005 :** *Statistique et probabilité, rappel de cours, QCM et question de réflexion, exercices corrigés, sujets d'Annales.* 3<sup>ème</sup> éditions. Dunod. Paris. 215p

**Calot G., 1980 :** *Cours de calcul des probabilités.* Dunod Décision. Paris. 476p

## ANNEXES

### I. Liste des tables des lois usuelles

**A1** : Probabilité individuelles de la loi de poisson  $P(\lambda)$

**A2** : Probabilité individuelles et cumulées de la loi de poisson  $P(\lambda)$

**A3** : Probabilité individuelles et cumulées de la loi de poisson  $P(\lambda)$

**A** : Fonction de répartition de la loi normale centrée et réduite  $\mathcal{N}(0, 1)$

**B** : Probabilité individuelles de la loi du Chi-deux  $\chi_v^2$

**C1** : Probabilité individuelles et cumulées de la loi Student Fisher  $t_{(v, \alpha)}$

**C2** : Probabilité individuelles et cumulées de la loi Student Fisher  $t_{(v, \alpha)}$

**D** : la loi Fisher-Snedecor  $F_{(v, \alpha)}$

### II. Tables de nombre au Hasard

**E1** : Extrait d'une table de nombre au Hasard de Kendall et Babington Smith, table tirée de Christian Labrousse, Statistique, Tome 2, Paris 1962



=====

ANNEXE A1 - Probabilités individuelles de la loi de Poisson  $\mathcal{P}(\lambda)$ .

Cette table donne  $p(\{X = k\}) = \frac{e^{-\lambda}\lambda^k}{k!}$  pour  $X \rightsquigarrow \mathcal{P}(\lambda)$  :

$k \backslash \lambda$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494
4		0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111
5				0,0001	0,0002	0,0004	0,0007	0,0012	0,0020
6							0,0001	0,0002	0,0003

$k \backslash \lambda$	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
0	0,3679	0,2231	0,1353	0,0821	0,0492	0,0302	0,0183	0,0111	0,0067
1	0,3679	0,3347	0,2707	0,2052	0,1454	0,1057	0,0733	0,0500	0,0337
2	0,1839	0,2510	0,2707	0,2565	0,2240	0,1850	0,1465	0,1125	0,0842
3	0,0613	0,1255	0,1804	0,2132	0,2240	0,2158	0,1954	0,1898	0,1404
4	0,0153	0,0471	0,0902	0,1336	0,1680	0,1888	0,1954	0,1898	0,1755
5	0,0031	0,0141	0,0361	0,0668	0,1008	0,1322	0,1563	0,1708	0,1755
6	0,0005	0,0035	0,0120	0,0278	0,0504	0,0771	0,1042	0,1281	0,1462
7	0,0001	0,0008	0,0034	0,0099	0,0216	0,0385	0,0595	0,0824	0,1044
8		0,0001	0,0009	0,0031	0,0081	0,0169	0,0298	0,0463	0,0653
9			0,0002	0,0009	0,0027	0,0066	0,0132	0,0232	0,0363
10				0,0002	0,0008	0,0023	0,0053	0,0104	0,0181
11					0,0002	0,0007	0,0019	0,0043	0,0082
12					0,0001	0,0002	0,0006	0,0016	0,0034
13						0,0001	0,0002	0,0006	0,0013
14							0,0001	0,0002	0,0005
15								0,0001	0,0002
16									0,0001

ANNEXE A2 - Probabilités individuelles et cumulées de la loi de Poisson  $\mathcal{P}(\lambda)$ .

Cette table donne  $p(\{X = k\}) = \frac{e^{-\lambda}\lambda^k}{k!}$  pour  $X \rightsquigarrow \mathcal{P}(\lambda)$  et  $F(k) = \sum_{l=0}^k e^{-\lambda}\frac{\lambda^l}{l!}$  :

$\lambda$	1		2		3		4		5	
$k$	$p(k, \lambda)$	$F(k)$								
0	0,3679	0,3679	0,1353	0,1353	0,0498	0,0498	0,0183	0,0183	0,0067	0,0067
1	0,3679	0,7358	0,2707	0,4060	0,1494	0,1991	0,0733	0,0916	0,0337	0,0404
2	0,1839	0,9197	0,2707	0,6767	0,2240	0,4232	0,1465	0,2381	0,0842	0,1247
3	0,0613	0,9810	0,1804	0,8571	0,2240	0,6472	0,1954	0,4335	0,1404	0,2650
4	0,0153	0,9963	0,0902	0,9473	0,1680	0,8152	0,1954	0,6288	0,1755	0,4405
5	0,0031	0,9994	0,0361	0,9834	0,1008	0,9161	0,1563	0,7851	0,1755	0,6160
6	0,0005	0,9999	0,0120	0,9955	0,0504	0,9665	0,1042	0,8893	0,1462	0,7622
7	0,0001	1,0000	0,0034	0,9989	0,0216	0,9881	0,0595	0,9489	0,1044	0,8666
8			0,0009	0,9998	0,0081	0,9962	0,0298	0,9786	0,0653	0,9319
9			0,0002	1,0000	0,0027	0,9989	0,0132	0,9919	0,0363	0,9682
10					0,0008	0,9997	0,0053	0,9972	0,0181	0,9863
11					0,0002	0,9999	0,0019	0,9991	0,0082	0,9945
12					0,0001	1,0000	0,0006	0,9997	0,0036	0,9980
13							0,0002	0,9999	0,0013	0,9993
14							0,0001	1,0000	0,0005	0,9998
15									0,0002	0,9999
16									0,0001	1,0000
$\lambda$	6		7		8		9		10	
$k$	$p(k, \lambda)$	$F(k)$								
0	0,0025	0,0025	0,0009	0,0009	0,0003	0,0003	0,0001	0,0001		
1	0,0149	0,0174	0,0064	0,0073	0,0027	0,0030	0,0011	0,0012	0,0005	0,0005
2	0,0446	0,0620	0,0223	0,0296	0,0107	0,0138	0,0050	0,0062	0,0023	0,0028
3	0,0892	0,1512	0,0521	0,0818	0,0286	0,0424	0,0150	0,0212	0,0076	0,0104
4	0,1339	0,2851	0,0912	0,1730	0,0573	0,0996	0,0337	0,0550	0,0189	0,0293
5	0,1606	0,4457	0,1277	0,3007	0,0916	0,1912	0,0607	0,1157	0,0378	0,0671
6	0,1606	0,6063	0,1490	0,4497	0,1221	0,3134	0,0911	0,2068	0,0631	0,1302
7	0,1377	0,7440	0,1490	0,5987	0,1396	0,4530	0,1171	0,3239	0,0901	0,2203
8	0,1033	0,8472	0,1304	0,7291	0,1396	0,5925	0,1318	0,4557	0,1126	0,3329
9	0,0688	0,9161	0,1014	0,8305	0,1241	0,7166	0,1318	0,5874	0,1251	0,4580
10	0,0413	0,9574	0,0710	0,9015	0,0993	0,8159	0,1186	0,7060	0,1251	0,5381
11	0,0225	0,9799	0,0452	0,9466	0,0722	0,8881	0,0970	0,8030	0,1137	0,6368
12	0,0113	0,9912	0,0264	0,9730	0,0481	0,9362	0,0728	0,8758	0,0948	0,7916
13	0,0052	0,9964	0,0142	0,9872	0,0296	0,9658	0,0504	0,9261	0,0729	0,8645
14	0,0022	0,9986	0,0071	0,9943	0,0169	0,9827	0,0324	0,9585	0,0521	0,9166
15	0,0009	0,9995	0,0033	0,9976	0,0090	0,9918	0,0194	0,9780	0,0347	0,9513
16	0,0003	0,9998	0,0014	0,9990	0,0045	0,9963	0,0109	0,9889	0,0217	0,9730
17	0,0001	1,0000	0,0006	0,9996	0,0021	0,9984	0,0058	0,9947	0,0128	0,9857
18			0,0002	0,9999	0,0009	0,9993	0,0029	0,9976	0,0071	0,9928
19			0,0001	1,0000	0,0004	0,9997	0,0014	0,9989	0,0037	0,9965
20					0,0002	0,9999	0,0006	0,9996	0,0019	0,9984
21					0,0001	1,0000	0,0003	0,9998	0,0009	0,9993
22							0,0001	0,9999	0,0004	0,9997
23								1,0000	0,0002	0,9999
24									0,0001	1,0000

ANNEXE A3 - Probabilités individuelles et cumulées de la loi de Poisson  $\mathcal{P}(\lambda)$ .

Cette table donne  $p(\{X = k\}) = \frac{e^{-\lambda} \lambda^k}{k!}$  pour  $X \rightsquigarrow \mathcal{P}(\lambda)$  et  $F(k) = \sum_{l=0}^k e^{-\lambda} \frac{\lambda^l}{l!}$  :

$\lambda$	11		12		13		14		15	
$k$	$p(k, \lambda)$	$F(k)$								
0										
1	0,0002	0,0002	0,0001	0,0001						
2	0,0010	0,0012	0,0004	0,0005	0,0002	0,0001	0,0001			
3	0,0037	0,0049	0,0018	0,0023	0,0008	0,0010	0,0004	0,0005	0,0002	0,0002
4	0,0102	0,0151	0,0053	0,0076	0,0027	0,0037	0,0013	0,0018	0,0007	0,0009
5	0,0224	0,0375	0,0127	0,0203	0,0070	0,0107	0,0037	0,0055	0,0019	0,0028
6	0,0411	0,0786	0,0255	0,0458	0,0152	0,0259	0,0087	0,0142	0,0048	0,0076
7	0,0646	0,1432	0,0437	0,0895	0,0281	0,0540	0,0174	0,0316	0,0104	0,0180
8	0,0888	0,2320	0,0655	0,1550	0,0457	0,0997	0,0304	0,0620	0,0194	0,0374
9	0,1085	0,3405	0,0874	0,2424	0,0661	0,1658	0,0473	0,1093	0,0324	0,0698
10	0,1194	0,4599	0,1048	0,3472	0,0859	0,2517	0,0663	0,1756	0,0486	0,1184
11	0,1194	0,5793	0,1144	0,4616	0,1015	0,3532	0,0844	0,2600	0,0663	0,1847
12	0,1094	0,6887	0,1144	0,5760	0,1099	0,4631	0,0984	0,3584	0,0829	0,2676
13	0,0926	0,7813	0,1056	0,6816	0,1099	0,5730	0,1060	0,4644	0,0956	0,3622
14	0,0728	0,8541	0,0905	0,7721	0,1021	0,6751	0,1060	0,5704	0,1024	0,4656
15	0,0534	0,9075	0,0724	0,8445	0,0885	0,7636	0,0989	0,6693	0,1024	0,5680
16	0,0367	0,9442	0,0543	0,8988	0,0719	0,8355	0,0866	0,7559	0,0960	0,6640
17	0,0237	0,9679	0,0383	0,9371	0,0550	0,8905	0,0713	0,8272	0,0847	0,7487
18	0,0145	0,9824	0,0255	0,9626	0,0397	0,9302	0,0554	0,8826	0,0706	0,8193
19	0,0084	0,9908	0,0161	0,9787	0,0272	0,9574	0,0409	0,9235	0,0558	0,8751
20	0,0046	0,9954	0,0097	0,9884	0,0177	0,9751	0,0286	0,9521	0,0418	0,9169
21	0,0024	0,9978	0,0055	0,9939	0,0109	0,9680	0,0191	0,9712	0,0299	0,9468
22	0,0012	0,9990	0,0030	0,9969	0,0065	0,9925	0,0121	0,9833	0,0204	0,9672
23	0,0006	0,9996	0,0016	0,9985	0,0037	0,9962	0,0074	0,9907	0,0133	0,9805
24	0,0003	0,9999	0,0008	0,9993	0,0020	0,9982	0,0043	0,9950	0,0083	0,9888
25	0,0001	1,0000	0,0004	0,9997	0,0010	0,9992	0,0024	0,9974	0,0050	0,9938
26			0,0002	0,9999	0,0005	0,9997	0,0013	0,9987	0,0029	0,9967
27			0,0001	1,0000	0,0002	0,9999	0,0007	0,9994	0,0016	0,9983
28					0,0001	1,0000	0,0003	0,9997	0,0009	0,9992
29							0,0002	0,9999	0,0004	0,9996
30							0,0001	1,0000	0,0002	0,9998
31									0,0001	0,9999
32									0,0001	1,0000

ANNEXE A - Fonction de répartition de la loi normale centrée réduite  $\mathcal{N}(0, 1)$ .

Cette table donne  $\Pi(x) = P\{X \leq x\}$  pour  $X \sim \mathcal{N}(0, 1)$  :

$x$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

ANNEXE B - Probabilités individuelles de la loi du  $\chi^2_\nu$ ,

Cette table donne les valeurs (quantiles)  $\chi^2_{\nu,1-\alpha}$  telles que  $p(\{\chi^2_\nu < \chi^2_{\nu,1-\alpha}\}) = 1 - \alpha$  :

$\nu \backslash 1 - \alpha$	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995
1	0,0000393	0,000157	0,000982	0,00393	0,0158	2,71	3,84	5,02	6,63	7,88
2	0,0100	0,0201	0,0506	0,103	0,211	4,61	5,99	7,38	9,21	10,60
3	0,072	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34	12,84
4	0,207	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,412	0,554	0,831	1,145	1,61	9,24	11,07	12,83	15,09	16,75
6	0,676	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,989	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	5,70	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
21	8,03	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40
22	8,64	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80
23	9,26	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18
24	9,89	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
25	10,52	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93
26	11,16	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64	48,29
27	11,81	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96	49,64
28	12,46	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28	50,99
29	13,12	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59	52,34
30	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67
40	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77
50	27,99	29,71	32,36	34,76	37,69	63,17	67,50	71,42	76,15	79,49
60	35,53	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38	91,95
70	43,28	45,44	48,76	51,74	55,33	85,53	90,53	95,02	100,4	104,2
80	51,17	53,54	57,15	60,39	64,28	96,58	101,9	106,6	112,3	116,3
90	59,20	61,75	65,65	69,13	73,29	107,6	113,1	118,1	124,1	128,3
100	67,33	70,06	74,22	77,93	82,36	118,5	124,3	129,6	135,8	140,2

ANNEXE C1 - Probabilités individuelles et cumulées de la loi de Student-Fischer  $t_{\nu, \alpha}$ ,

Cette table donne les valeurs (quantiles)  $t_{\nu, 1-\alpha}$  telles que  $p(\{-t_{\nu} < t_{\nu, 1-\alpha}\}) = 1 - \alpha$  :

$\nu \backslash 1 - \alpha$	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	0,975	0,99	0,995	0,9995
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	635,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,150
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,648
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
$\infty$	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

ANNEXE C2 - Probabilités individuelles et cumulées de la loi de Student-Fischer  $t_{\nu,\alpha}$ .

Cette table donne les valeurs  $t_{\nu,\alpha}$  telles que  $p(\{t_{\nu,\alpha} < t_\nu < +t_{\nu,\alpha}\}) = 1 - \alpha$  :

$\nu \backslash \alpha$	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,387	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,745	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,649
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,656
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
$\infty$	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,940	2,326	2,576	3,291

ANNEXE D - la loi de Fischer-Snedecor,

Cette table donne, pour  $\alpha = 0,025$ , pour  $\nu_1$  et  $\nu_2$  donnés, les valeurs  $F_{\nu_1, \nu_2, 1-\alpha}$  telles que  $p(\{X < F_{\nu_1, \nu_2, 1-\alpha}\}) = 1 - \alpha$ ,

$\nu_2 \backslash \nu_1$	$\infty$	120	60	40	30	24	20	15	12	10	9	8	7	6	5	4	3	2	1
1	1018	1014	1010	1006	1001	997,2	993,1	984,9	976,7	968,6	963,3	956,7	948,2	937,1	921,8	899,6	864,2	794,6	617,6
2	39,50	39,49	39,48	39,47	39,46	39,45	39,45	39,43	39,41	39,40	39,39	39,37	39,36	39,33	39,30	39,25	39,17	39,00	38,51
3	13,90	13,95	13,99	14,04	14,08	14,12	14,17	14,26	14,34	14,42	14,47	14,54	14,62	14,73	14,88	15,10	15,44	16,04	17,44
4	8,26	8,31	8,36	8,41	8,46	8,51	8,56	8,66	8,75	8,84	8,90	8,98	9,07	9,20	9,36	9,60	9,98	10,65	12,22
5	6,02	6,07	6,12	6,18	6,23	6,28	6,33	6,43	6,52	6,62	6,68	6,76	6,86	6,98	7,16	7,39	7,76	8,43	10,01
6	4,85	4,90	4,96	5,01	5,07	5,12	5,17	5,27	5,37	5,46	5,52	5,60	5,70	5,82	5,99	6,23	6,60	7,26	8,81
7	4,14	4,20	4,25	4,31	4,36	4,42	4,47	4,57	4,67	4,76	4,82	4,90	4,99	5,12	5,29	5,52	5,89	6,54	8,07
8	3,67	3,73	3,78	3,84	3,89	3,95	4,00	4,10	4,20	4,30	4,36	4,43	4,53	4,65	4,82	5,05	5,42	6,06	7,57
9	3,33	3,39	3,45	3,51	3,56	3,61	3,67	3,77	3,87	3,96	4,03	4,10	4,20	4,32	4,48	4,72	5,08	5,71	7,21
10	3,08	3,14	3,20	3,26	3,31	3,37	3,42	3,52	3,62	3,72	3,78	3,85	3,95	4,07	4,24	4,47	4,83	5,46	6,94
11	2,88	2,94	3,00	3,06	3,12	3,17	3,23	3,33	3,43	3,53	3,59	3,66	3,76	3,88	4,04	4,28	4,63	5,26	6,72
12	2,72	2,79	2,85	2,91	2,96	3,02	3,07	3,18	3,28	3,37	3,44	3,51	3,61	3,73	3,89	4,12	4,47	5,10	6,55
13	2,60	2,66	2,72	2,78	2,84	2,89	2,95	3,05	3,15	3,25	3,31	3,39	3,48	3,60	3,77	4,00	4,35	4,97	6,41
14	2,49	2,55	2,61	2,67	2,73	2,79	2,84	2,95	3,05	3,15	3,21	3,29	3,38	3,50	3,66	3,90	4,24	4,86	6,30
15	2,40	2,46	2,52	2,58	2,64	2,70	2,76	2,86	2,96	3,06	3,12	3,20	3,29	3,41	3,58	3,82	4,15	4,77	6,20
16	2,32	2,38	2,45	2,51	2,57	2,63	2,68	2,79	2,89	2,99	3,05	3,12	3,22	3,34	3,50	3,73	4,08	4,69	6,12
17	2,25	2,32	2,38	2,44	2,50	2,56	2,62	2,72	2,82	2,92	2,98	3,06	3,16	3,28	3,44	3,66	4,01	4,62	6,04
18	2,19	2,26	2,32	2,38	2,44	2,50	2,56	2,67	2,77	2,87	2,93	3,01	3,10	3,22	3,38	3,61	3,95	4,56	5,96
19	2,13	2,20	2,27	2,33	2,39	2,45	2,51	2,62	2,72	2,82	2,88	2,96	3,05	3,17	3,33	3,56	3,90	4,51	5,90
20	2,09	2,16	2,22	2,28	2,35	2,41	2,46	2,57	2,68	2,77	2,84	2,91	2,99	3,11	3,29	3,52	3,86	4,46	5,87
21	2,04	2,11	2,18	2,25	2,31	2,37	2,42	2,53	2,64	2,73	2,80	2,87	2,97	3,09	3,25	3,48	3,82	4,42	5,83
22	2,00	2,08	2,14	2,21	2,27	2,33	2,39	2,50	2,60	2,70	2,76	2,84	2,93	3,05	3,22	3,44	3,78	4,38	5,79
23	1,97	2,04	2,11	2,18	2,24	2,30	2,36	2,47	2,57	2,67	2,73	2,81	2,90	3,02	3,18	3,41	3,75	4,35	5,75
24	1,94	2,01	2,08	2,15	2,21	2,27	2,33	2,44	2,54	2,64	2,70	2,78	2,87	2,99	3,15	3,38	3,72	4,32	5,72
25	1,91	1,98	2,05	2,12	2,18	2,24	2,30	2,41	2,51	2,61	2,68	2,75	2,86	2,97	3,13	3,35	3,69	4,29	5,69
26	1,88	1,95	2,03	2,09	2,16	2,22	2,28	2,39	2,49	2,59	2,65	2,73	2,82	2,94	3,10	3,33	3,67	4,27	5,66
27	1,85	1,93	2,00	2,07	2,13	2,19	2,25	2,36	2,47	2,57	2,63	2,71	2,80	2,92	3,08	3,31	3,65	4,24	5,63
28	1,83	1,91	1,98	2,05	2,11	2,17	2,23	2,34	2,45	2,55	2,61	2,69	2,78	2,90	3,06	3,29	3,63	4,22	5,61
29	1,81	1,89	1,96	2,03	2,09	2,15	2,21	2,32	2,43	2,53	2,59	2,67	2,76	2,88	3,04	3,27	3,61	4,20	5,59
30	1,79	1,87	1,94	2,01	2,07	2,14	2,20	2,31	2,41	2,51	2,57	2,65	2,75	2,87	3,03	3,25	3,59	4,18	5,57
40	1,64	1,72	1,80	1,88	1,94	2,01	2,07	2,18	2,28	2,39	2,45	2,53	2,62	2,74	2,90	3,13	3,46	4,05	5,42
60	1,48	1,57	1,65	1,74	1,81	1,88	1,94	2,06	2,17	2,27	2,33	2,41	2,50	2,63	2,79	3,01	3,34	3,93	5,29
120	1,31	1,41	1,51	1,61	1,69	1,76	1,82	1,94	2,05	2,16	2,22	2,30	2,39	2,52	2,67	2,89	3,23	3,80	5,15
$\infty$	1,00	1,27	1,39	1,48	1,57	1,64	1,71	1,83	1,94	2,05	2,11	2,19	2,29	2,41	2,57	2,79	3,12	3,69	5,02

## EXTRAITS D'UNE TABLE DE NOMBRES AU HASARD

(Kendall et Babington Smith, table tirée de Christian Labrousse, Statistique, Tome2, Dunod, Paris, 1962)

02 22 85 19 48 74 55 24 89 69 15 53 00 20 88 48 95 08  
 85 76 34 51 40 44 62 93 65 99 72 64 09 34 01 13 09 74  
 00 88 96 79 38 24 77 00 70 91 47 43 43 82 71 67 49 90  
 64 29 81 85 50 47 36 50 91 19 09 15 98 75 60 58 33 15  
 94 03 80 04 21 49 54 91 77 85 00 45 68 23 12 94 23 44  
 42 28 52 73 06 41 37 47 47 31 52 99 89 82 22 81 86 55  
 09 27 52 72 49 11 30 93 33 29 54 17 54 48 47 42 04 79  
 54 68 64 07 85 32 05 96 54 79 57 43 96 97 30 72 12 19  
 25 04 92 29 71 11 64 10 42 23 23 67 01 19 20 58 35 93  
 28 58 32 91 95 28 42 36 98 59 66 32 15 51 46 63 57 10  
 64 35 04 62 24 87 44 85 45 68 41 66 19 17 13 09 63 37  
 61 05 55 88 25 01 15 77 12 90 69 34 36 93 52 39 36 23

98 93 18 93 86 98 99 04 75 28 30 05 12 09 57 35 90 15  
 61 89 35 47 16 32 20 16 78 52 82 37 26 33 67 42 11 93  
 94 40 82 18 06 61 54 67 03 66 76 82 90 31 71 90 39 27  
 54 38 58 65 27 70 93 57 59 00 63 56 18 79 85 52 21 03  
 63 70 89 23 76 46 97 70 00 62 15 35 97 42 47 54 60 60  
 61 58 65 62 81 29 69 71 95 53 53 69 20 95 66 60 50 70  
 51 68 98 15 05 64 43 32 74 07 44 63 52 38 67 59 56 69  
 59 25 41 48 64 79 62 26 87 86 94 30 43 54 26 98 61 38  
 85 00 02 24 67 85 88 10 34 01 54 53 23 77 33 11 19 68  
 01 46 87 56 19 19 19 43 70 25 24 29 48 22 44 81 35 40  
 42 41 25 10 87 27 77 28 05 90 73 03 95 46 88 82 25 02  
 03 57 14 03 17 80 47 85 94 49 89 55 10 37 19 50 20 37  
 18 95 93 40 45 43 04 56 17 03 34 54 83 91 69 02 90 72