

Fondamentaux de **statistique** pour le **management**

- Exercices d'application
- Problèmes corrigés
- Rappels de cours



Roxane Cattan-Jallet
Lynn Farah
Corinne Hahn



Fondamentaux de **statistique** pour le **management**

Roxane CATTAN-JALLET

Professeur à ESSCA Business School & ESCP Business School

Lynn FARAH

Professeur à ESCP Business School

Corinne HAHN

Professeur à ESCP Business School



Conception graphique couverture : Nathalie FOULLOY
Conception graphique intérieur : Agathe GALLET DUCAROY
Mise en pages : Nord Compo

ISBN 9782340-117686

Dépôt légal : août 2026

©Ellipses Édition Marketing S.A.
8/10 rue la Quintinie 75015 Paris



Le Code de la propriété intellectuelle et artistique n'autorisant, aux termes des alinéas 2 et 3 de l'article L. 122-5, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause, est illicite » (alinéa 1^{er} de l'article L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

www.editions-ellipses.fr

Sommaire

Thème 1. Statistique Descriptive Univariée	7
Chapitre 1. Population, échantillon, variable statistique	9
Chapitre 2. Tableaux et graphiques	19
Chapitre 3. Indicateurs statistiques	35
 Thème 2. Statistique Descriptive Bivariée	 55
Chapitre 4. Dépendance statistique	57
Chapitre 5. Corrélation	71
Chapitre 6. Corrélation linéaire	81
Problème de synthèse (Chapitre 1 à 6) : Les usagers des réseaux sociaux	95
 Thème 3. Statistique Inférentielle	 101
Chapitre 7. Estimation	103
Chapitre 8. Test du khi-deux	115
Chapitre 9. Tests de conformité	129
Chapitre 10. Tests de comparaison sur échantillons indépendants	141
Problème de synthèse (Chapitre 7 à 10) : Les bienfaits du mariage	154
 Thème 4. Modélisation et prévision	 161
Chapitre 11. Séries chronologiques	163
Chapitre 12. Régression linéaire simple	185
Chapitre 13. Régression linéaire multiple	203
Chapitre 14. Analyse de la Variance (ANOVA)	223
Problème de synthèse (Chapitre 11 à 14) : Les logements Airbnb	236

Introduction

Ce livre propose un panorama des méthodes statistiques couramment utilisées en management. Il est ancré dans l'expérience d'enseignement des auteurs en Bachelor, Master et formation continue.

Ce livre se compose de 14 chapitres structurés en 4 thèmes : statistique descriptive univariée, statistique descriptive bivariée, statistique inférentielle, modélisation et prévision.

Chacun de ces chapitres se compose de 3 parties :

- Une synthèse des principales notions de cours, illustrées par un exemple dans le domaine du management ;
- Des exercices, issus de la pratique et de niveaux de difficultés croissants, portant sur les différents savoir-faire associés aux notions exposées dans le chapitre, savoir-faire qui sont précisés au début de chaque énoncé ;
- Des corrigés détaillés incluant des alertes sur les erreurs couramment rencontrées, des suggestions de méthodes alternatives et des points d'approfondissement.

Afin de mettre en musique les différentes notions abordées dans les chapitres d'un même thème, chaque thème se conclut par un problème managérial synthétique accompagné de son corrigé.

Ainsi les étudiants seront en mesure de traiter l'ensemble des notions indispensables à la pratique de leur futur métier. Ce livre leur permet également d'approfondir certaines notions en fonction de besoins spécifiques qui peuvent émerger au cours de leur formation ou de leur pratique professionnelle.

Thème 1

Statistique Descriptive Univariée

Population, échantillon, variable statistique

Introduction

Exemple : Black Friday

Le Black Friday constitue une occasion privilégiée pour observer les comportements d'achat liés aux grandes vagues de promotions.

Dans ce cadre, une enseigne mène une étude sur l'ensemble des 200 clients ayant effectué au moins un achat durant la journée du Black Friday en 2025 dans son enseigne. Pour chacun de ces clients, plusieurs informations ont été recueillies : le montant dépensé, la durée de fidélité, la satisfaction exprimée après l'achat, le niveau de revenu, le nombre d'articles achetés, la proportion de produits retournés, le niveau de réduction obtenu, ainsi que la catégorie de produits privilégiée et le mode d'achat, en ligne ou en magasin. L'analyse de ces données va permettre de dégager des tendances, de comparer les habitudes des consommateurs et de mieux comprendre l'impact réel du Black Friday sur les comportements d'achat.

La *statistique* permet de mener ce type d'étude. Cette discipline regroupe un ensemble de méthodes scientifiques dont l'objectif est de décrire, analyser, structurer et modéliser des informations.

Afin de collecter ces informations, il faut définir clairement la population sur laquelle porte l'étude et décider si on effectue une étude sur toute ou partie de la population.

Nous présentons dans ce chapitre les notions fondamentales de la statistique.

1. Concepts élémentaires

Il convient de définir en amont certains concepts élémentaires.

- La *population* désigne l'ensemble des unités, personnes ou objets, de l'étude.
- Un *individu* est une unité de la population.

- La *taille de la population*, notée N , correspond au nombre d'individus contenus dans la population.
- Un *caractère* est un critère observé sur les individus étudiés.
- Un *recensement* est un recueil d'observations qui porte sur l'ensemble des N individus de la population.
- Un *n-échantillon* est un groupe de n individus extrait de la population selon une méthode spécifique.

Exemple : Black Friday

La population est représentée par l'ensemble des 200 clients ayant effectué au moins un achat durant la journée du Black Friday dans l'enseigne étudiée. Pour chaque client, correspondant à un individu, neuf caractères sont observés : le montant dépensé (en euros), la durée de fidélité (en mois), la satisfaction exprimée après l'achat (note sur 10), le niveau de revenu (en euros), le nombre d'articles achetés, la proportion de produits retournés (en pourcentage), le niveau de réduction obtenu (en pourcentage), la catégorie de produits privilégiée (alimentation, électronique, maison ou vêtements) et le mode d'achat (en ligne ou en magasin). Dans cet exemple, nous prenons en compte l'ensemble des clients ayant effectué un achat le jour du Black Friday, c'est-à-dire la population sur laquelle porte l'étude.

2. Variable statistique

Lorsqu'à chaque individu correspond un choix unique par rapport au critère étudié, nous pouvons construire une *variable statistique* qui définira une partition de la population étudiée.

Il existe deux types de variables : les *variables qualitatives* et les *variables quantitatives*.

a. Variable quantitative

Définitions

- Une *variable quantitative* est associée à un caractère mesurable.
- On appelle *valeurs*, les réalisations possibles d'une variable quantitative.
- Une variable quantitative est dite *discrète* lorsque ses valeurs sont dénombrables (ou limitées).
- Si les valeurs de la variable quantitative ne sont pas dénombrables, alors la variable est dite quantitative *continue*.

Exemple : Black Friday

Parmi les neuf caractères observés sur les 200 clients, on distingue sept caractères mesurables :

- 1 variable quantitative discrète : le nombre d'articles achetés prenant 5 valeurs possibles (valeurs de 1 à 5) ;
- 6 variables quantitatives continues : le montant dépensé, la durée de fidélité, la satisfaction exprimée après l'achat, le niveau de revenu, la proportion de produits retournés et le niveau de réduction obtenu.

La satisfaction exprimée après l'achat, notée sur 10, est considérée comme une variable continue car elle peut prendre une valeur décimale. Par exemple, le premier client a attribué une note de 6,6/10.

b. Variable qualitative

Définitions

- Une *variable qualitative* est associée à un caractère non mesurable.
- On appelle *modalités*, les réalisations possibles d'une variable qualitative.
- Une variable qualitative est dite *ordinaire* lorsque ses modalités peuvent se ranger dans un ordre préétabli.
- Si les modalités de la variable qualitative ne peuvent pas être classées, alors la variable est dite qualitative *nominale*.

→ Remarques

- Les modalités prises par une variable *qualitative* peuvent être éventuellement identifiées par des nombres mais il ne s'agit pas pour autant d'une variable quantitative. En effet, effectuer des opérations arithmétiques sur ces nombres ne fait pas sens. C'est le cas par exemple pour les numéros des départements français.
- Les variables qualitatives ordinales sont souvent traitées comme des variables quantitatives lorsqu'elles comportent suffisamment de degrés et que ces degrés peuvent être considérés comme équidistants.

Par exemple, la satisfaction des clients d'un magasin peut être mesurée à l'aide des modalités ordonnées : --, -, 0, + et ++. Il est alors possible d'associer pour chacune de ces modalités une valeur numérique correspondante allant de 0 à 4.

Exemple : Black Friday

Parmi les neuf caractères observés sur les 200 clients, on distingue deux caractères non mesurables : la catégorie de produits privilégiée comportant quatre modalités (alimentation, électronique, maison ou vêtements), ainsi que le mode d'achat à deux modalités (en ligne ou en magasin).

Les modalités de ces variables ne pouvant pas être ordonnées, il s'agit de deux variables qualitatives nominales.

3. Méthodes d'échantillonnage

La constitution d'un n -échantillon à partir d'une population peut se faire selon différentes méthodes d'échantillonnage : les méthodes aléatoires de sondage et les méthodes empiriques d'échantillonnage.

a. Les méthodes aléatoires de sondage

Dans un *sondage aléatoire*, chaque individu a une probabilité connue et non nulle d'appartenir à l'échantillon. Un *échantillon aléatoire* est constitué par un mécanisme aléatoire qui respecte ces différentes probabilités.

→ Remarque

Dans le cas où les individus ont tous la même probabilité d'appartenir à la population, le modèle de référence s'appelle le *sondage aléatoire simple*.

On appelle *taux de sondage*, la taille de l'échantillon divisée par la taille de la population, soit n/N .

→ Remarques

- Le taux de sondage doit être faible. On considère généralement qu'il est satisfaisant lorsqu'il est inférieur à 10 %.
- Le sondage aléatoire simple est en fait peu utilisé dans la pratique, notamment car, dans de nombreux cas, la base de sondage n'est pas connue.

b. Les méthodes empiriques d'échantillonnage

Dans les *méthodes empiriques*, la sélection des données n'est pas effectuée par un processus aléatoire. La méthode la plus connue est la méthode des quotas.

Un échantillon construit par la *méthode des quotas* est un échantillon qui respecte la répartition de certaines caractéristiques au sein de la population. On dira qu'il est *représentatif* de la population étudiée parce qu'il représente certaines caractéristiques de cette population.

→ **Remarque**

Un échantillon représentatif n'est pas nécessairement de bonne qualité puisqu'il ne considère que certaines des caractéristiques de la population.

Exemple : Black Friday

L'étude a été réalisée auprès de 200 clients, qui constituent directement la population étudiée. Il ne s'agit donc pas d'un échantillon.

Exercice 1**Identifier la population et le caractère étudiés**

Dans un immeuble de bureaux, à chaque étage on a mis une machine à café à disposition des salariés. Pour chaque étage, on relève le nombre de cafés consommés par jour ainsi que le nombre de salariés travaillant à l'étage.

1. Définir la population sur laquelle porte l'étude.
2. Déterminer le(s) caractère(s) de la population étudié(s).

Exercice 2**Identifier population, échantillon et variable(s) à partir d'un tableau**

Une chaîne d'agences de voyage désire étudier la qualité de ses prestations de service dans les trois agences qui font le plus gros chiffre d'affaires. Pour ce faire, elle fait établir par un bureau d'étude la distribution des arrivées de la clientèle à ses guichets. On note alors x_i , le nombre de personnes entrant dans ses agences pour chacune des 2 000 périodes de 2 minutes choisies aléatoirement dans les heures les plus chargées de la journée.

Nous obtenons le tableau suivant :

x_i	0	1	2	3	4	5	6	7	8	9
n_i	163	418	500	437	270	128	57	19	6	2

1. Définir la population sur laquelle porte l'étude ainsi que l'échantillon qui a été sélectionné.
2. Déterminer la variable considérée dans cette étude en précisant sa nature.

Exercice 3**Identifier population, échantillon et variable(s) à partir d'un tableau**

Une société d'assurance possède des agences réparties sur le territoire français. Le tableau ci-après donne pour les 11 premières agences, le nombre d'employés dans l'agence ainsi que le nombre de contrats d'assurance vendus par mois.

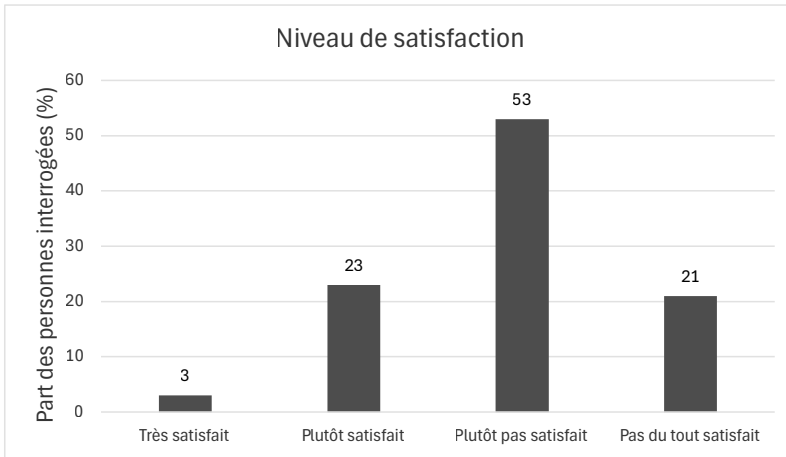
Nombre d'employés dans l'agence	Nombre de contrats vendus par mois
10	21
12	42
28	45
38	64
30	38
22	30
36	66
44	66
52	78
8	24
24	16

1. Définir la population sur laquelle porte l'étude ainsi que l'échantillon qui a été sélectionné.
2. Déterminer la ou les variables considérées dans cette étude en précisant leur nature.

Exercice 4

Identifier population, échantillon et variable(s) à partir d'un graphique

On demande à 500 étudiants français choisis aléatoirement, s'ils sont satisfaits des réponses des candidats à l'élection présidentielle aux préoccupations des étudiants.



1. Définir la population sur laquelle porte l'étude ainsi que l'échantillon qui a été sélectionné.
2. Au regard du graphique, déterminer la ou les variables considérées dans cette étude en précisant leur nature.

Exercice 5

Identifier la nature des variables

La Direction des Ressources Humaines d'une entreprise consigne dans un tableau les indicateurs suivants sur l'ensemble de ses salariés :

- Le salaire (en k€)
- L'année d'embauche
- L'ancienneté (en mois)
- Le nombre d'enfants à charge
- Le statut (0 si junior et 1 si sénior)

Pour chacune de ces variables, préciser s'il s'agit de variables qualitatives nominales/ordinales ou de variables quantitatives discrètes/continues.

Exercice 1

1. La population correspond à l'ensemble des étages observés.

Un individu correspond à un étage et la population correspond donc à l'ensemble des étages.

△ L'étude menée ici est un recensement puisqu'elle porte sur l'ensemble de la population.

2. On étudie ici, deux variables : le nombre de cafés consommés par jour et le nombre de salariés travaillant à l'étage.

Exercice 2

1. La population est l'ensemble des périodes de 2 minutes.

L'échantillon sélectionné correspond à 2 000 périodes de 2 minutes.

2. La variable étudiée est le nombre de personnes entrant dans les agences. Il s'agit d'une variable quantitative discrète prenant 10 valeurs distinctes (de 0 à 9).

Exercice 3

1. La population est l'ensemble des agences.

L'échantillon est composé des 11 premières agences.

△ L'échantillon n'est pas aléatoire puisqu'il est constitué des 11 premières agences.

2. Les variables étudiées sont :

- Le nombre d'employés dans l'agence,
- Le nombre de contrats vendus par l'agence.

Ces deux variables sont des variables quantitatives discrètes.

△ Dans le cas de variables discrètes présentant de nombreuses valeurs hétérogènes, on choisit en général de regrouper ces valeurs dans des classes et donc de traiter ces variables comme des variables continues.

Exercice 4

1. La population correspond à l'ensemble des étudiants.

L'échantillon est composé des 500 étudiants sélectionnés aléatoirement.

2. La variable étudiée est la satisfaction. Il s'agit d'une variable qualitative ordinale à 4 modalités : Très satisfait ; Plutôt satisfait ; Plutôt pas satisfait ; Pas du tout satisfait.

Exercice 5

- Le salaire (en k€) : Variable quantitative continue
- L'année d'embauche : Variable qualitative ordinale
- L'ancienneté (en mois) : Variable quantitative continue
- Le nombre d'enfants à charge : Variable quantitative discrète
- Le statut (0 si junior et 1 si senior) : Variable qualitative nominale

△ L'année d'embauche, bien qu'exprimée par un nombre, n'est pas une variable quantitative (la somme de deux années n'a aucun sens). De la même manière, les 0 et 1 de la variable statut sont le résultat d'un codage. Ces valeurs bien que numériques sont des modalités.

Tableaux et graphiques

Introduction

Exemple : Black Friday

Lors de l'opération Black Friday, les informations recueillies auprès des 200 clients apparaissent sous la forme d'une liste brute de données dont la lecture n'est pas aisée.

Voici un extrait du tableau de données :

Consommateur	Montant dépensé (€)	Fidélité (mois)	Satisfaction (1-10)	Salaire (€)	Nombre d'articles	Taux de retour (%)	Taux de réduction (%)	Catégorie principale	Moyen d'achat
1	108,00	35,0	6,6	2021,83	3	33,33	26,85	Vêtements	En ligne
2	179,65	38,8	7,8	2694,98	3	0,00	42,71	Maison	En magasin
3	79,40	11,8	2,4	1729,03	1	100,00	2,47	Alimentation	En ligne
4	64,28	46,4	8,9	1520,45	4	0,00	41,67	Alimentation	En ligne
5	246,64	45,5	8,8	3469,85	5	0,00	42,20	Electronique	En ligne
6	211,85	3,2	1,3	3041,46	5	20,00	14,25	Electronique	En ligne
7	170,22	45,1	8,4	2725,55	2	0,00	23,14	Electronique	En ligne
8	163,13	13,6	2,2	2475,72	2	50,00	21,93	Maison	En ligne
9	60,38	18,8	4,0	1636,89	5	20,00	10,93	Alimentation	En ligne
10	111,81	39,2	7,7	2201,42	5	0,00	19,84	Alimentation	En magasin

Pour pouvoir étudier les comportements d'achat, il est nécessaire de disposer d'outils permettant de faciliter la compréhension des données recueillies.

Après la collecte et avant l'analyse, on cherche à structurer et représenter les données recueillies.

Ce chapitre est consacré à la présentation des principaux tableaux et graphiques permettant la description des données.

1. Tableaux

a. Distribution d'une variable

Définitions

- La *distribution* d'une variable est résumée dans un tableau appelé *tableau de distribution*.
- Pour une variable qualitative (nominale ou ordinale) ou pour une variable quantitative discrète, le tableau de distribution présente les *effectifs* observés pour chaque modalité ou valeur prise par la variable, c'est-à-dire le nombre d'individus qui prennent cette modalité ou cette valeur.
- Pour une variable quantitative continue, le tableau de distribution présente les *effectifs* observés pour chaque classe de valeurs prises par la variable, c'est-à-dire le nombre d'individus qui appartiennent à cette classe.

→ Remarque

Il est plus simple de constituer des classes de même *amplitude* en créant des intervalles de même longueur. Toutefois, selon la distribution, il est souvent nécessaire d'envisager des classes d'amplitudes inégales.

Exemple : Black Friday

Prenons la variable qualitative correspondant au mode d'achat, le tableau de distribution est :

Mode d'achat	Effectifs
En ligne	102
En magasin	98
Total	200

On constate que parmi les 200 clients, la répartition entre achat en ligne et en magasin est très équilibrée.

Prenons la variable quantitative discrète correspondant au nombre d'articles achetés, le tableau de distribution est :

Nombre d'articles achetés	Effectifs
1	34
2	38
3	39
4	48
5	41
Total	200

On constate que les effectifs associés aux différentes valeurs sont très proches, ce qui traduit une répartition uniforme du nombre d'articles achetés.

Prenons la variable quantitative continue correspondant au montant dépensé, le tableau de distribution est :

Montant dépensé	Effectifs
[0;75[	15
[75;150[	79
[150;225[	87
[225;300[	19
Total	200

Nous avons constitué quatre classes disjointes de même amplitude (75).

On constate que les effectifs observés varient fortement d'une classe à l'autre. Il est souvent plus pertinent de répartir les données différemment en affinant les classes là où les observations sont particulièrement concentrées.