

# Table des matières

## Se repérer dans l'ouvrage

### **Avant-propos**

<i>Les objectifs et définitions incontournables</i> .....	3
---	---

### **Introduction**

<i>Les données</i> .....	7
--------------------------	---

I. Quel type de données ? .....	7
I.1. Définitions : recensement et sondage, données individuelles et agrégées.....	7
I.2. Où trouver des données ? .....	10
II. Les données issues d'enquêtes par sondage .....	12
II.1. Les principales méthodes d'échantillonnage et leur biais intrinsèque : nous avons affaire à des échantillons !.....	12
II.2. Le « répondant fantôme » : un deuxième biais de mesure de l'enquête par questionnaire.....	20
II.3. La désirabilité sociale : un troisième biais de l'enquête par questionnaire....	21
II.4. Comment rectifier l'échantillonnage <i>a posteriori</i> ? .....	22
III. Les exercices d'application .....	25
III.1. Logement et travaux d'isolation .....	25
III.2. Le vote Front national .....	26
III.3. Les corrigés .....	27
IV. Les articles d'application conseillés .....	29

### **Chapitre 1.**

<i>Le point de départ à l'analyse statistique : le fichier de données et ses variables</i> .....	31
--	----

I. Le vocabulaire du fichier de données .....	34
I.1. Questions, batteries d'items, variables .....	34
I.2. Questions fermée, ouverte, semi-ouverte, modalités, valeurs, verbatims.....	37
I.3. Questions à choix unique versus à choix multiple .....	40
II. Le codage des données .....	41
II.1. Le précodage.....	41
II.2. Le post-codage.....	45

III.	Les types de variables.....	46
III.1.	Les variables nominales : un premier cas particulier de variable qualitative .....	47
III.2.	Les variables ordinaires : un second cas particulier de variable qualitative .....	48
III.3.	Les variables quantitatives.....	49
III.4.	Les cas limites : les variables dichotomiques et les variables ordinaires comprenant un nombre important de modalités.....	49
III.5.	Rendez-vous avec ces différents types de variables dans les chapitres à venir ! .....	51
IV.	La distribution statistique d'une variable .....	51
IV.1.	La distribution sous forme de tableau .....	52
IV.2.	La distribution sous forme de représentation graphique .....	54
V.	Les exercices d'application .....	57
V.1.	L'enquête <i>Vin et pesticides</i> : typage et précodage des questions .....	57
V.2.	L'enquête <i>Vin et pesticides</i> : distribution statistique des variables .....	57
V.3.	Les corrigés .....	58
VI.	Les articles d'application conseillés .....	59

## **Chapitre 2**

*Décrire une variable quantitative :*

les indicateurs de tendance centrale et de dispersion .....	61	
I.	Les indicateurs de tendance centrale .....	63
I.1.	Le(s) mode(s).....	63
I.2.	La moyenne.....	65
I.3.	La médiane.....	69
I.4.	De la nécessité d'aller au-delà des indicateurs de tendance centrale .....	73
II.	Les indicateurs de dispersion .....	76
II.1.	La dispersion autour de la moyenne : l'écart-type.....	76
II.2.	La dispersion autour de la médiane : quantiles, écart et rapport interquartiles.....	79
III.	Les exercices d'application .....	86
III.1.	La comparaison de durées d'interview.....	86
III.2.	Le taux d'abstention en Gironde au second tour de l'élection présidentielle de 2012.....	87
III.3.	Les corrigés .....	87
IV.	Les articles d'application conseillés .....	90

### **Chapitre 3**

<i>La préparation des variables à l'analyse statistique : du recodage à la construction d'indicateurs.....</i>	91
I. Le recodage d'une variable.....	92
I.1. Précautions préalables à prendre .....	92
I.2. Deux difficultés fréquentes : les faibles effectifs et la présence de données manquantes .....	93
I.3. Comment recoder : une démarche hypothético-déductive ou inductive ? ...	96
II. La création d'indicateurs à partir de plusieurs variables .....	103
II.1. Les patrons de réponses .....	107
II.2. Les compteurs d'occurrence .....	109
II.3. Les échelles d'attitude.....	112
III. Études de cas.....	117
III.1. Les patrons de réponses .....	117
III.2. Les compteurs d'occurrence .....	123
III.3. Les échelles d'attitude.....	126
IV. Les exercices d'application .....	129
IV.1. Impact de l'UE dans les domaines de l'emploi et de la politique sociale.....	129
IV.2. Pour aller plus loin : revenus par UC et dotation patrimoniale .....	130
IV.3. Les corrigés .....	131
V. Les articles d'application conseillés .....	136

### **Chapitre 4**

<i>L'inférence statistique .....</i>	137
I. Le passage des résultats dans l'échantillon à la conclusion sur la population d'intérêt : l'inférence statistique.....	137
I.1. Inférence et intervalle de confiance .....	137
I.2. Une brève introduction aux tests statistiques : la logique du raisonnement .....	143
II. Les exercices d'application .....	144
II.1. La lecture de la table de distribution de la loi normale centrée réduite ....	144
II.2. Le calcul d'intervalles de confiance .....	145
II.3. Pour aller plus loin : la comparaison de sondages.....	146
II.4. Les corrigés .....	146
III. Les articles d'application conseillés .....	148

## **Chapitre 5**

<i>La mesure du lien statistique entre deux variables quantitatives .....</i>	149
I. Le nuage de points .....	150
I.1. La lecture d'un nuage de points à partir d'exemples.....	150
I.2. La forme et la dispersion d'un nuage de points.....	152
II. La mesure statistique : le coefficient de corrélation linéaire .....	153
II.1. La formule : le R de Pearson.....	153
II.2. Les caractéristiques du coefficient de corrélation linéaire .....	154
II.3. L'interprétation du coefficient de corrélation linéaire à partir d'exemples .....	155
III. Les limites du coefficient de corrélation linéaire .....	156
III.1. La linéarité .....	156
III.2. L'analyse des points aberrants ( <i>outliers</i> ) .....	157
III.3. L'erreur écologique .....	158
IV. Les exercices d'application .....	162
IV.1. Le taux de chômage chez les femmes et chez les hommes .....	162
IV.2. La densité de médecins et de lits d'hôpitaux pour 1 000 habitants .....	162
IV.3. Le tableau de corrélations : Europe 2020.....	163
IV.4. Pour aller plus loin : l'analyse par sous-groupes .....	164
IV.5. Les corrigés .....	165
V. Les articles d'application conseillés .....	168

## **Chapitre 6**

<i>La mesure du lien statistique entre deux variables nominales ou ordinaires .....</i>	169
I. La présentation des tableaux croisés .....	169
I.1. Le tableau des effectifs .....	170
I.2. Les tableaux des pourcentages.....	171
II. La comparaison des pourcentages .....	174
II.1. Les écarts à la moyenne.....	175
II.2. Les rapports .....	176
II.3. Les rapports de chance ( <i>odds ratios</i> ) .....	177
III. La mesure de la significativité de la relation : le test du Chi-deux .....	178
III.1. Étape 1 : poser l'hypothèse que l'on cherche à tester .....	179
III.2. Étape 2 : calculer la statistique de test correspondante (c'est-à-dire une valeur chiffrée unique).....	179
III.3. Étape 3 : conclure sur le test statistique .....	183
III.4. La procédure complète : existe-t-il un lien statistiquement significatif entre l'indicateur revenus/patrimoine et l'opinion à propos de l'item	

« Pour établir la justice sociale, il faudrait prendre aux riches pour donner aux pauvres » ? .....	186
IV. La mesure de l'intensité de la relation .....	186
IV.1. Les limites du Chi-deux .....	186
IV.2. Le V de Cramer.....	187
V. Les exercices d'application .....	192
V.1. La lecture de la table de la distribution de la loi du Chi-deux.....	192
V.2. L'opinion sur le fonctionnement de la démocratie .....	193
V.3. Les corrigés .....	193
VI. Les articles d'application conseillés .....	195
<b>Chapitre 7</b>	
<i>La mesure du lien statistique entre une variable explicative qualitative et une variable à expliquer quantitative .....</i>	197
I. Les principes de la comparaison de moyenne .....	197
I.1. L'analyse descriptive .....	197
I.2. La question de la variance .....	198
II. Le cas d'une variable explicative qualitative à deux modalités : le test de Student ou le test normal ?.....	199
II.1. La comparaison des variances : le test de Fisher .....	200
II.2. La comparaison des moyennes : test de Student ou test normal .....	201
II.3. Un exemple de procédure complète : la comparaison des durées d'interviews dans la vague 5 de l' <i>European Social Survey</i> en France et en Suède .....	206
III. Le cas d'une variable explicative qualitative à plus de deux modalités : l'analyse de variance .....	207
III.1. La décomposition de la variance.....	207
III.2. L'ANOVA .....	208
III.3. Un exemple de procédure complète : la comparaison des durées d'interviews dans la vague 5 de l' <i>European Social Survey</i> en France, en Espagne et en Suède.....	211
IV. Les exercices d'application .....	214
IV.1. La lecture des tables des distributions de la loi de Student et de celle de Fisher .....	214
IV.2. L'âge moyen à la décohabitation selon le sexe .....	216
IV.3. La durée d'interview selon le diplôme.....	216
IV.4. Les corrigés .....	217
V. Les articles d'application conseillés .....	220

<b>Chapitre 8</b>	<b>221</b>
<i>Les représentations graphiques .....</i>	221
I. Quelle représentation graphique ? .....	221
I.1. Les questions à se poser pour trancher .....	221
I.2. Un exemple pas à pas à partir d'une variable à expliquer nominale : passage de l'analyse univariée à la multivariée .....	225
II. Les règles fondamentales de la représentation graphique.....	229
II.1. Des critères typographiques parfois imposés.....	229
II.2. La graduation des axes : un choix qui n'est pas neutre .....	230
II.3. L'homogénéité des axes pour une comparaison rigoureuse.....	232
III. Les cas particuliers : les données spatio-temporelles .....	236
III.1. La représentation d'une évolution temporelle .....	236
III.2. La représentation des données géographiques .....	239
IV. Les exercices d'application .....	247
IV.1. Le souhait de devenir propriétaire.....	247
IV.2. Les qualités et défauts des candidats .....	247
IV.3. La densité de population et le type de logement .....	248
IV.4. Pour aller plus loin : le taux de croissance du PIB .....	249
IV.5. Les corrigés .....	249
V. Les articles d'application conseillés .....	253
<b>Bibliographie</b>	
<i>Compléments de lecture .....</i>	255
Les articles et ouvrages d'application.....	255
Les manuels en méthodes d'analyse statistique.....	260