

Table des matières

1	Éléments de la statistique inférentielle	13
1.1	Introduction	13
1.2	Types de variables	13
1.3	Éléments d'échantillonnage	14
1.3.1	Population étudiée	14
1.3.2	Base de sondage	15
1.3.3	Techniques d'échantillonnage	15
1.3.4	Taille de l'échantillon	17
1.3.5	Mise en œuvre	17
1.4	Statistique descriptive et statistique inférentielle	17
1.5	Liaisons entre variables	18
1.6	Tests d'hypothèse	18
1.6.1	p-value associée à un test	20
1.6.2	Mise en œuvre d'un test	20
1.6.3	Catégories de tests	20
1.6.4	Mesure de l'association entre deux variables qualitatives	21
1.6.5	Mesure de l'association entre deux variables quantitatives	21
1.6.6	Variabes quantitatives/qualitatives	21
	Test d'indépendance entre deux variables qualitatives	23
2.1	Introduction	23
2.2	Données et notations	23
2.3	Exemple	25
2.4	Autres distributions	26
2.4.1	Distributions marginales	26
2.4.2	Distribution des fréquences	27
2.4.3	Distributions conditionnelles	27
2.5	Le modèle statistique	29
2.5.1	Tableau des effectifs théoriques	30

2.5.2	Indicateur du Khi-deux	31
2.5.3	Propriétés	31
2.5.4	Test du Khi-deux	33
2.5.5	Contributions au χ^2	35
2.6	Mesures d'association	35
2.6.1	Coefficient phi	36
2.6.2	Coefficient de contingence	36
2.6.3	V de Cramer	36
2.6.4	T de Tschuprow	37
2.7	Limites du test	38
2.7.1	Effectifs théoriques trop faibles	38
2.7.2	Variables cachées	39
2.8	Graphiques	40
2.8.1	Diagramme en barres	40
2.8.2	Diagramme en barres empilées	40
2.8.3	Diagramme en mosaïques	41
2.8.4	Diagramme d'écart à l'indépendance	43
2.9	Compléments	44
2.9.1	Cas des tableaux 2×2	45
2.9.2	Test de McNemar	46
2.9.3	Analyse Facorielle des Correspondances	48
2.10	Utilisation du logiciel R	50
2.10.1	Création d'un tableau de contingence	50
2.10.2	Autres tableaux	50
2.10.3	Test du Khi-deux	51
2.10.4	Graphiques	52
2.10.5	Test de McNemar	53
2.10.6	Mesures d'association	53
2.10.7	Analyse des correspondances simples	54
2.11	Résumons-nous	54
2.12	Pour aller plus loin	55
2.13	Exercices	55
	La régression linéaire simple	59
3.1	Introduction	59
3.2	Données et notations	60
3.3	Exemple	60
3.4	Aspects descriptifs	60
3.4.1	Représentation graphique : le nuage de points	61
3.4.2	Droite des moindres carrés	61
3.4.3	Propriétés	65
3.4.4	Formule de décomposition de la variation	67
3.4.5	Coefficient de détermination	69

3.5	Aspects inférentiels	69
3.5.1	Le modèle statistique	69
3.5.2	Estimation des coefficients de régression	70
3.5.3	Intervalle de confiance de la pente b	71
3.5.4	Intervalle de confiance de l'intercept a	72
3.5.5	Tests sur les paramètres	73
3.5.6	Intervalle de confiance pour la valeur moyenne de Y	73
3.5.7	Intervalle de prédiction pour une valeur de Y	74
3.5.8	Test F	75
3.5.9	Tableau de l'analyse de la variance de la régression	77
3.6	Diagnostics	78
3.6.1	Hypothèses du modèle	78
3.6.2	Linéarité	78
3.6.3	Indépendance des erreurs	80
3.6.4	Égalité des variances	80
3.6.5	Normalité des erreurs	80
3.6.6	Détection des observations atypiques	82
3.7	Compléments	82
3.7.1	Démonstrations	83
3.7.2	Régression multiple	86
3.7.3	Régression non linéaire	87
3.8	Utilisation de R	89
3.8.1	Régression linéaire simple	89
3.8.2	Régression linéaire multiple	94
3.8.3	Régression non-linéaire	96
3.9	Résumons-nous	96
3.10	Pour aller plus loin	96
3.11	Exercices	96

Test d'indépendance entre deux variables quantitatives 101

4.1	Introduction	101
4.2	Données et notations	101
4.3	Exemple	102
4.4	Aspects descriptifs	103
4.4.1	Représentation graphique	103
4.4.2	Covariance	103
4.4.3	Coefficient de corrélation linéaire	105
4.5	Aspects inférentiels	109
4.5.1	Cadre statistique	110
4.5.2	Le test	110
4.5.3	Emploi de la table du coefficient de corrélation	111
4.6	Corrélation et causalité	112
4.7	Compléments	112

4.7.1	Loi de T	112
4.7.2	Intervalle de confiance d'un coefficient de corrélation	113
4.7.3	Matrice des corrélations	114
4.7.4	Corrélation multiple	115
4.7.5	Corrélation partielle	117
4.8	Utilisation de \mathbf{R}	120
4.8.1	Corrélation linéaire simple	120
4.8.2	Intervalle de confiance de r	121
4.8.3	Matrice des corrélations	122
4.8.4	Corrélation multiple	122
4.8.5	Corrélation partielle	122
4.9	Résumons-nous	123
4.10	Pour aller plus loin	123
4.11	Exercices	123
5	Liaison entre deux variables ordinales	127
5.1	Introduction	127
5.2	Données et notations	128
5.3	Exemple	128
5.4	Principe	129
5.5	Coefficient de corrélation de Spearman	130
5.5.1	Définition	130
5.5.2	Le test	131
5.6	Coefficient de corrélation de Kendall	132
5.6.1	Définition	132
5.6.2	Calculs Pratiques	134
5.6.3	Le test	135
5.7	Cas des variables quantitatives	136
5.8	Compléments	138
5.8.1	Présence d'ex-æquos	138
5.8.2	Corrélation de rang partielle de Kendall	138
5.8.3	Concordance de p classements	140
5.9	Utilisation de \mathbf{R}	141
5.9.1	Corrélation de Spearman	141
5.9.2	Corrélation de Kendall	142
5.9.3	Variables quantitatives	142
5.10	Résumons-nous	143
5.11	Pour aller plus loin	143
5.12	Exercices	143

6	Comparaison de deux populations	147
6.1	Introduction	147
6.2	Données et notations	148
6.3	Exemple	148
6.4	Représentations graphiques	149
6.5	Test de comparaison de deux moyennes	150
6.5.1	Cadre statistique	150
6.5.2	Test t de Student	151
6.5.3	Application et décision	152
6.6	Test d'égalité de deux variances	153
6.6.1	Statistique du test	153
6.6.2	Application et décision	154
6.7	Compléments	155
6.7.1	Comparaison de deux proportions	155
6.7.2	Test de Wilcoxon (Mann-Whitney)	157
6.7.3	Échantillons appariés	159
6.8	Utilisation de R	160
6.8.1	Représentations des données	160
6.8.2	Test d'égalité de deux variances	161
6.8.3	Test d'égalité de deux moyennes	162
6.8.4	Comparaison de deux proportions	162
6.8.5	Test de Wilcoxon (Mann-Whitney)	163
6.8.6	Échantillons appariés	163
6.9	Résumons-nous	164
6.10	Pour aller plus loin	165
6.11	Exercices	165
7	Analyse de la variance à un facteur	169
7.1	Introduction	169
7.2	Données et notations	169
7.3	Exemple	170
7.4	Aspects descriptifs	171
7.4.1	Représentation des données	171
7.4.2	Formule de décomposition de la variation	173
7.4.3	Rapport de corrélation	175
7.5	Aspects inférentiels	176
7.5.1	Cadre statistique	176
7.5.2	Le test de Fisher	176
7.5.3	Tableau d'analyse de la variance	177
7.5.4	Comparaisons multiples	178
7.6	Compléments	180
7.6.1	Test de Bartlett	180
7.6.2	Test de Levene	181

7.6.3	Comparaison de plusieurs proportions	182
7.6.4	Test de Kruskal-Wallis	183
7.6.5	Analyse de la variance à plusieurs facteurs	184
7.6.6	Analyse de la covariance	186
7.7	Utilisation de R	188
7.7.1	Analyse de la variance	188
7.7.2	Comparaisons multiples	189
7.7.3	Test de Bartlett	190
7.7.4	Test de Levene	190
7.7.5	Comparaison de plusieurs proportions	191
7.7.6	Test de Kruskal-Wallis	191
7.7.7	Analyse de la variance à plusieurs facteurs	192
7.7.8	Analyse de la covariance	192
7.8	Résumons-nous	193
7.9	Pour aller plus loin	193
7.10	Exercices	193
8	Études de Cas	197
8.1	Introduction	197
8.2	Étude 1	197
8.3	Étude 2	211
9	Corrigé des exercices	225
9.1	Chapitre 2	225
9.2	Chapitre 3	229
9.3	Chapitre 4	233
9.4	Chapitre 5	238
9.5	Chapitre 6	240
9.6	Chapitre 7	244
A	Symbole \sum	249
A.1	Introduction	249
A.2	Définition	249
A.3	Propriétés	250
A.4	Double somme	250
A.5	Exercices	251
A.6	Corrigé	252
B	Lois de Probabilités	253
B.1	Introduction	253
B.2	Loi de Bernoulli	253
B.3	Loi binomiale $\mathcal{B}(n, \pi)$	253
B.4	Loi multinomiale	254
B.4.1	Définition et propriétés	254

B.4.2	Application fondamentale	255
B.5	Loi normale	256
B.5.1	Définition et propriétés	256
B.5.2	Adéquation à une loi normale	257
B.6	Loi du Khi-deux	259
B.6.1	Définition et propriétés	259
B.6.2	Application fondamentale	261
B.7	Loi de Student	261
B.7.1	Définition et propriétés	261
B.7.2	Application fondamentale	263
B.8	Loi de Fisher	264
B.8.1	Définition et propriétés	264
B.8.2	Application fondamentale	265
C	Logiciel R	267
C.1	Introduction	267
C.2	Démarrage	267
C.3	Opérateurs élémentaires	268
C.4	Vecteurs et matrices	268
C.5	Les fonctions classiques	270
C.6	Quelques fonctions de probabilités	270
C.7	Graphiques	272
C.8	Tests classiques	272
D	Tables statistiques	275