


Marie-Noëlle Jubénot
Daniel Eudes

ANALYSE DES DONNÉES SOUS R

pour LES SCIENCES HUMAINES

Théories et exemples commentés

Package DataBoo
à télécharger



ellipses



Table des matières

Introduction	5
Partie 1. Analyse descriptive et notions de base	11
Chapitre 1. Analyse descriptive sur les variables quantitatives	12
I.1.1. Les principaux indicateurs du centre d'une variable.....	12
I.1.2. Variables quantitatives : les concepts de variance, covariance et de corrélation	18
I.1.3. Variables quantitatives : les matrices de variance-covariance et de corrélation.....	33
Chapitre 2. Critères de liaisons entre deux variables qualitatives	62
I.2.1. Construction du tableau de contingence et des matrices des profils	62
I.2.2. Tester la liaison entre deux variables qualitatives : test du Khi-deux, le V de Cramér et le T de Tschuprow	66
Chapitre 3. Les calculs de proximités entre individus	73
I.3.1. Définition d'un indice de distance ou de proximité.....	73
I.3.2. La distance euclidienne.....	74
I.3.3. La distance de Gower.....	75
Partie 2. Introduction à R et à l'interface RStudio et présentation de la base de données	77
Chapitre 1. Installation et présentation des logiciels R et RStudio	79
II.1.1. Installation du logiciel R	79
II.1.2. Installation de l'interface RStudio	83
II.1.3. Présentation de RStudio	84
Chapitre 2. Initiation au langage R	112
II.2.1. Le concept d'« objet ».....	112
II.2.2. Les fonctions de base du logiciel R.....	131
II.2.3. L'utilisation des packages	143

Partie 3. Méthodes d'analyse factorielle	147
Chapitre 1. Analyse en Composantes Principales, normée et non normée	148
III.1.1. Diagonalisation de la matrice de corrélation	149
III.1.2. Critères d'interprétation des résultats et représentations graphiques.....	157
III.1.3. ACP : exercices d'application avec R.....	178
Chapitre 2. L'Analyse Factorielle des Correspondances	188
III.2.1. Calcul des distances et de l'inertie	188
III.2.2. Mise en œuvre de l'Analyse Factorielle des Correspondances.....	200
III.2.3. Critères d'interprétation et d'analyse	207
III.2.4. AFC : exercices d'application avec R.....	209
Chapitre 3. Analyse des Correspondances Multiples	213
III.3.1. Le Tableau Disjonctif Complet généralisé	213
III.3.2. Propriétés du tableau de Burt.....	218
III.3.3. Mise en œuvre de l'ACM	221
III.3.4. Critères d'interprétation de l'ACM.....	222
III.3.5. ACM : exercices d'application avec R.....	224
Chapitre 4. Analyse simultanée de données quantitatives et qualitatives	226
III.4.1. Tester la dépendance entre une variable quantitative et une variable qualitative...226	
III.4.2. Tester la dépendance entre plusieurs variables quantitatives et qualitatives.....	230
III.4.3. AFDM : exercices d'application avec R.....	234
Partie 4. Méthodes de classification	241
Chapitre 1. Les concepts de base	243
IV.1.1. Ensembles, partitions, hiérarchies.....	243
IV.1.2. Interprétation d'un dendrogramme.....	245
IV.1.3. Les distances couramment utilisées en classification	247
Chapitre 2. Classification Ascendante Hiérarchique (CAH)	251
IV.2.1. Inertie inter, inertie intra et théorème de Huygens	251
IV.2.2. L'algorithme de la CAH.....	253
IV.2.3. Choix du nombre de classes.....	263
IV.2.4. CAH : exercices d'application avec R.....	265
Chapitre 3. Méthode de partitionnement <i>K-means</i>	269
IV.3.1. L'algorithme <i>K-means</i>	269
IV.3.2. Le choix du nombre de classes	272
IV.3.3. <i>K-means</i> : exercices d'application avec R.....	274
ANNEXES	277
Annexe 1. Liste des tableaux	278
Annexe 2. Liste des graphiques	279
Annexe 3. Liste des encadrés	279
Annexe 4. Liste des scripts R	280