

Revue de Statistique Appliquée -
Mars 2006.

BIBLIOGRAPHIE

MODÈLES STATISTIQUES POUR DONNÉES QUALITATIVES

par J.J. DROESBEKE, M. LEJEUNE, G. SAPORTA

1 vol, 308 pages, Éditions TECHNIP (27 rue Ginoux, 75737, Paris, cedex 15) 2005,
48 euros, ISBN 2 7108-0855-2

Cet ouvrage, qui présente d'une part les méthodes exploratoires et factorielles permettant de décrire les liaisons entre variables qualitatives, et d'autre part les modèles permettant d'expliquer une variable qualitative en fonction de facteurs qualitatifs ou quantitatifs, fait suite aux neuvièmes journées d'études en statistique organisées par la SFdS en 2000. Fruit de la collaboration de 11 auteurs, il comporte les 12 chapitres suivants :

Quelques aspects historiques du traitement des données qualitatives – les coefficients d'association et les tests d'indépendance pour des variables qualitatives – le modèle log-linéaire – les classes latentes – modèles linéaires généralisés – modèles à réponse dichotomique-comparaison de méthodes de discrimination – régression logistique robuste – modèles de comptage et applications en assurance – mesures répétées – le modèle à réponse multinominale, application à la lutte contre le goître endémique – la régression logistique PLS.

Une bibliographie d'environ 200 références et un index terminent cet ouvrage très clair et agréable à lire. De nombreux exemples illustrent les méthodes exposées.

Ce manuel constitue un excellent cours sur la modélisation des données catégorielles, avec les développements les plus récents sur le sujet. Il s'adresse aux statisticiens, aux étudiants de Master, et de façon plus générale aux praticiens devant modéliser des variables qualitatives.

DATA MINING ET STATISTIQUE DÉCISIONNELLE

L'intelligence dans les bases de données

par S. TUFFÉRY

1 vol, 398 pages, Editions TECHNIP (27 rue Ginoux, 75737, Paris, cedex 15) 2005,
45 euros, ISBN 2 7108-0867-6

Cet ouvrage, traite du data mining qui est l'application des techniques de statistique, d'analyse des données et d'intelligence artificielle à l'exploration et à l'analyse de grandes bases de données en vue d'en extraire les informations utiles et pertinentes pour le décideur. Il comporte 15 chapitres. Les cinq premiers sont relatifs à une introduction sur le data mining avec en particulier le déroulement d'une

étude de data mining, l'exploration et la préparation des données, l'utilisation de données commerciales, et un aperçu sur les techniques de data mining. Les chapitres 6 à 10 sont consacrés aux méthodes de data mining : analyse factorielle, réseaux de neurones, classification automatique, techniques de classement et de prédiction, recherche de règles d'association etc. Les derniers chapitres traitent d'applications diverses (scoring, text mining, Web mining), des logiciels de statistique et de data mining ainsi que des facteurs de succès d'un projet de data mining. Deux annexes (rappels de statistique et data mining, informatique et libertés) et une intéressante bibliographie commentée terminent ce manuel, très clair et très complet.

Cet ouvrage de référence où les questions techniques sont abordées avec la rigueur voulue, mais sans formalisme excessif, et qui comporte de nombreux exemples d'application traités avec les logiciels SAS, SPAD et SPSS s'adresse aux statisticiens, aux utilisateurs et gestionnaires de bases de données, aux décideurs aux économètres, aux enseignants et aux étudiants en sciences économiques, aux « data miners » etc.

STATISTIQUES GÉNÉRALES POUR UTILISATEURS 1-MÉTHODOLOGIE

par J. PAGÈS

1 vol, 218 pages, Presse Universitaires de Rennes (UHB, Rennes 2,
Campus de la Harpe, 2 rue du Doyen Denis Leroy, 35044, Rennes, Cedex) 2005,
14 euros, ISBN 2-7535-0164-5

Ce livre, qui est une transcription d'un cours de statistique donné aux étudiants d'Agrocampus Rennes s'appuie sur l'exemple pour exposer les principales méthodes statistiques utiles aux praticiens : statistique descriptive, estimation, tests de comparaison de moyennes et de variances, analyse de variance, régression simple et multiple, test du chi-2, plans d'expérience, analyse en composantes principales.

Pour chaque technique développée, un petit problème concret permet de bien comprendre la méthodologie statistique utilisée pour le résoudre, avec les calculs effectués, et l'interprétation des résultats.

Au niveau pédagogique, l'auteur va du particulier au général, ce qui permet au lecteur de progresser pas à pas dans la connaissance.

Ce livre d'introduction à la statistique pratique très complet, et qui ne comporte pratiquement pas de démonstrations, ni d'aspects formels généraux concernant l'inférence statistique constitue un excellent manuel très clair d'apprentissage à la statistique. Il s'adresse aux étudiants en écoles d'agronomie (la plupart des exemples traités sont issus du domaine agronomique) et aussi à tout étudiant désirant s'initier par l'exemple à la pratique statistique.