1. Record Nr. UNINA9910975000303321

Autore	Upton Graham J. G
Titolo	Categorical Data Analysis by Example
Pubbl/distr/stampa	Newark : , : John Wiley & Sons, Incorporated, , 2016 ©2017
ISBN	9781119307914 9781119307860
Edizione	[1st ed.]
Descrizione fisica	1 online resource (215 pages)
Disciplina	519.535
Soggetti	Multivariate analysis
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- CATEGORICAL DATA ANALYSIS BY EXAMPLE -- Contents -- Preface -- Acknowledgments -- 1 Introduction -- 1.1 What are Categorical Data? -- 1.2 A Typical Data Set -- 1.3 Visualization and Cross-Tabulation -- 1.4 Samples, Populations, and Random Variation -- 1.5 Proportion, Probability, and Conditional Probability -- 1.6 Probability Distributions -- 1.6.1 The Binomial Distribution -- 1.6.2 The Multinomial Distribution -- 1.6.3 The Poisson Distribution -- 1.6.4 The Normal Distribution -- 1.6.5 The Chi-Squared (2) Distribution -- 1.7 *The Likelihood -- 2 Estimation and Inference for Categorical Data -- 2.1 Goodness of Fit -- 2.1.1 Pearson's X2 Goodness-of-Fit Statistic -- 2.1.2 *The Link between X2 and the Poisson and 2-Distributions -- 2.1.3 The Likelihood-Ratio Goodness-of-Fit Statistic, G2 -- 2.1.4 *Why the G2 and X2 Statistics Usually have Similar Values -- 2.2 Hypothesis Tests for a Binomial Proportion (Large Sample) -- 2.2.1 The Normal Score Test -- 2.2.2 *Link to Pearson's X2 Goodness-of-Fit Test -- 2.2.3 G2 for a Binomial Proportion -- 2.3 Hypothesis Tests for a Binomial Proportion (Small Sample) -- 2.3.1 One-Tailed Hypothesis Test -- 2.3.2 Two-Tailed Hypothesis Tests -- 2.4 Interval Estimates for a Binomial Proportion -- 2.4.1 Laplace's Method -- 2.4.2 Wilson's Method -- 2.4.3 The Agresti-Coull Method -- 2.4.4 Small Samples and Exact Calculations -- References -- 3 The 2 × 2 Contingency Table -- 3.1 Introduction -- 3.2 Fisher's Exact Test (for Independence) -- 3.2.1

*Derivation of the Exact Test Formula -- 3.3 Testing Independence with Large Cell Frequencies -- 3.3.1 Using Pearson's Goodness-of-Fit Test -- 3.3.2 The Yates Correction -- 3.4 The 2 × 2 Table in a Medical Context -- 3.5 Measuring Lack of Independence (Comparing Proportions) -- 3.5.1 Difference of Proportions -- 3.5.2 Relative Risk -- 3.5.3 Odds-Ratio -- References.
4 The I × J Contingency Table -- 4.1 Notation -- 4.2 Independence in the I × J Contingency Table -- 4.2.1 Estimation and Degrees of Freedom -- 4.2.2 Odds-Ratios and Independence -- 4.2.3 Goodness of Fit and Lack of Fit of the Independence Model -- 4.3 Partitioning -- 4.3.1 *Additivity of G2 -- 4.3.2 Rules for Partitioning -- 4.4 Graphical Displays -- 4.4.1 Mosaic Plots -- 4.4.2 Cobweb Diagrams -- 4.5 Testing Independence with Ordinal Variables -- References -- 5 The Exponential Family -- 5.1 Introduction -- 5.2 The Exponential Family -- 5.2.1 The Exponential Dispersion Family -- 5.3 Components of a General Linear Model -- 5.4 Estimation -- References -- 6 A Model Taxonomy -- 6.1 Underlying Questions -- 6.1.1 Which Variables are of Interest? -- 6.1.2 What Categories should be Used? -- 6.1.3 What is the Type of Each Variable? -- 6.1.4 What is the Nature of Each Variable? -- 6.2 Identifying the Type of Model -- 7 The 2 × J Contingency Table -- 7.1 A Problem with X2 (and G2) -- 7.2 Using the Logit -- 7.2.1 Estimation of the Logit -- 7.2.2 The Null Model -- 7.3 Individual Data and Grouped Data -- 7.4 Precision, Confidence Intervals, and Prediction Intervals -- 7.4.1 Prediction Intervals -- 7.5 Logistic Regression with a Categorical Explanatory Variable -- 7.5.1 Parameter Estimates with Categorical Variables (J &gt -- 2) -- 7.5.2 The Dummy Variable Representation of a Categorical Variable -- References -- 8 Logistic Regression with Several Explanatory Variables -- 8.1 Degrees of Freedom when there are no Interactions -- 8.2 Getting a Feel for the Data -- 8.3 Models with two-Variable Interactions -- 8.3.1 Link to the Testing of Independence between Two Variables -- 9 Model Selection and Diagnostics -- 9.1 Introduction -- 9.1.1 Ockham's Razor -- 9.2 Notation for Interactions and for Models -- 9.3 Stepwise Methods for Model Selection Using G2.
9.3.1 Forward Selection -- 9.3.2 Backward Elimination -- 9.3.3 Complete Stepwise -- 9.4 AIC and Related Measures -- 9.5 The Problem Caused by Rare Combinations of Events -- 9.5.1 Tackling the Problem -- 9.6 Simplicity Versus Accuracy -- 9.7 DFBETAS -- References -- 10 Multinomial Logistic Regression -- 10.1 A Single Continuous Explanatory Variable -- 10.2 Nominal Categorical Explanatory Variables -- 10.3 Models for an Ordinal Response Variable -- 10.3.1 Cumulative Logits -- 10.3.2 Proportional Odds Models -- 10.3.3 Adjacent-Category Logit Models -- 10.3.4 Continuation-Ratio Logit Models -- References -- 11 Log-Linear Models for I × J Tables -- 11.1 The Saturated Model -- 11.1.1 Cornered Constraints -- 11.1.2 Centered Constraints -- 11.2 The Independence Model for an I × J Table -- 12 Log-Linear Models for I × J × K Tables -- 12.1 Mutual Independence: ABC -- 12.2 The Model ABC -- 12.3 Conditional Independence and Independence -- 12.4 The Model ABAC -- 12.5 The Models ABACBC and ABC -- 12.6 Simpson's Paradox -- 12.7 Connection between Log-Linear Models and Logistic Regression -- Reference -- 13 Implications and Uses of Birch's Result -- 13.1 Birch's Result -- 13.2 Iterative Scaling -- 13.3 The Hierarchy Constraint -- 13.4 Inclusion of the All-Factor Interaction -- 13.5 Mostellerizing -- References -- 14 Model Selection for Log-Linear Models -- 14.1 Three Variables -- 14.2 More than Three Variables -- Reference -- 15 Incomplete Tables, Dummy Variables, and Outliers -- 15.1 Incomplete Tables -- 15.1.1 Degrees of Freedom -- 15.2 Quasi-independence --

15.3 Dummy Variables -- 15.4 Detection of Outliers -- 16 Panel Data and Repeated Measures -- 16.1 The Mover-Stayer Model -- 16.2 The Loyalty Model -- 16.3 Symmetry -- 16.4 Quasi-Symmetry -- 16.5 The Loyalty-Distance Model -- References -- Appendix R Code for Cobweb Function -- Index -- Author Index.
Index of Examples -- EULA.

Sommario/riassunto

Introduces the key concepts in the analysis of categoricaldata with illustrative examples and accompanying R code This book is aimed at all those who wish to discover how to analyze categorical data without getting immersed in complicated mathematics and without needing to wade through a large amount of prose. It is aimed at researchers with their own data ready to be analyzed and at students who would like an approachable alternative view of the subject. Each new topic in categorical data analysis is illustrated with an example that readers can apply to their own sets of data. In many cases, R code is given and excerpts from the resulting output are presented. In the context of log-linear models for cross-tabulations, two specialties of the house have been included: the use of cobweb diagrams to get visual information concerning significant interactions, and a procedure for detecting outlier category combinations. The R code used for these is available and may be freely adapted. In addition, this book: Uses an example to illustrate each new topic in categorical data Provides a clear explanation of an important subject Is understandable to most readers with minimal statistical and mathematical backgrounds Contains examples that are accompanied by R code and resulting output Includes starred sections that provide more background details for interested readers Categorical Data Analysis by Example is a reference for students in statistics and researchers in other disciplines, especially the social sciences, who use categorical data. This book is also a reference for practitioners in market research, medicine, and other fields. GRAHAM J. G. UPTON is formerly Professor of Applied Statistics, Department of Mathematical Sciences, University of Essex. Dr. Upton is author of The Analysis of Cross-tabulated Data (1978) and joint author of Spatial Data Analysis by Example (2 volumes, 1995), both published by Wiley. He is the lead author of The Oxford Dictionary of Statistics (OUP, 2014). His books have been translated into Japanese, Russian, and Welsh. "