Modeling Count Data

Deftly balancing theory and application, this book stands out in its coverage of the derivation of the GLM families and their foremost links. This edition has new sections on discrete response models, including zero-truncated, zero-inflated, censored, and hurdle count models, as well as heterogeneous negative binomial, and more.
Modeling Count DataCambridge University Press
This entry-level text offers clear and concise guidelines on how to select, construct, interpret, and evaluate count data. Written for researchers with little or no background in advanced statistics, the book presents treatments of all major models using numerous tables, insets, and detailed modeling suggestions. It begins by demonstrating the fundamentals of modeling count data, including a thorough present
model. It then works up to an analysis of the problem of overdispersion and of the negative binomial model, and finally to the many variations that can be made to the base count models. Examples in Stata, R, and SAS code enable readers to adapt models for their own purposes, making the text an ideal resource for researchers working in health, ecology, econometrics, transportation, and other fields.
Learn Poisson and negative binomial regression techniques About This Video Work with single and multiple independent variables Learn about model visualization and probability prediction Get acquainted with the dataset and explore continuous and binary variables In Detail The course is divided into two parts. In the first part, you'll be introduced to the theory behind count models in an intuitive way while math at a minimum. The course starts with an overview of count tables, where you'll learn how to calculate the incidence rate ratio. You'll get to grips with Poisson regression and understand how to work with continuous, binary, and categorical variables. As you advance, you'll explore the concept of overdispersion and how to address this issue using negative binomial models. The course also covers other as truncated models and zero-inflated models. In the second part of the course, you'll be able to apply what you have learned using Stata. You'll be taken through a large project where you'll fit the Poisson, negative binomial, and zero-inflated models. Additionally, you'll discover the tools used to compare these models.
Probability and Bayesian Modeling
Bayesian Modeling Using WinBUGS
Econometric Analysis of Count Data
Best Practices and Modern Methods
International Encyclopedia of Statistical Science
Linear Modeling for Unbalanced Data, Second Edition

From the creator of the popular website Ask a Manager and New York's work-advice columnist comes a witty, practical guide to 200 difficult professional conversations—featuring all-new advice! There's a reason Alison Green has been called "the Dear Abby of the work world." Ten years as a workplace-advice columnist have taught her that people avoid awkward conversations in the office because they simply don't know what to say. Thankfully, Green does—and in this incredibly helpful book, she tackles the tough discussions you may need to have during your career. You'll learn what to say when • coworkers push their work on you—then take credit for it • you accidentally trash-talk someone in an email then hit "reply all" • you're being micromanaged—or not being managed at all • you catch a colleague in a lie • your boss seems unhappy with your work • your cubemate's loud speakerphone is making you homicidal • you got drunk at the holiday party Praise for Ask a Manager "A must-read for anyone who works . . . [Alison Green's] advice boils down to the idea that you should be professional (even when others are not) and that communicating in a straightforward manner with candor and kindness will get you far, no matter where you work."—Booklist (starred review) "The author's friendly, warm, no-nonsense writing is a pleasure to read, and her advice can be widely applied to relationships in all areas of readers' lives. Ideal for anyone new to the job market or new to management, or anyone hoping to improve their work experience."—Library Journal (starred review) "I am a huge fan of Alison Green's Ask a Manager column. This book is even better. It teaches us how to deal with many of the most vexing big and little problems in our workplaces—and to do so with grace, confidence, and a sense of humor."—Robert Sutton, Stanford professor and author of The No Asshole Rule and The Asshole Survival Guide "Ask a Manager is the ultimate playbook for navigating the traditional workforce in a diplomatic but firm way."—Erin Lowry, author of Broke Millennial: Stop Scraping By and Get Your Financial Life Together
The goal of this book is multidimensional: a) to help reviving Statistics education in many parts in the world where it is in crisis. For the first time authors from many developing countries have an opportunity to write together with the most prominent world authorities. The editor has spent several years searching for the most reputable statisticians all over the world. International contributors are either presidents of the local statistical societies, or head of the Statistics department at the main university, or the most distinguished statisticians in their countries. b) to enable any non-statistician to obtain quick and yet comprehensive and highly understandable view on certain statistical term, method or application c) to enable all the researchers, managers and practicioners to refresh their knowledge in Statistics, especially in certain controversial fields. d) to revive interest in statistics among students, since they will see its usefulness and relevance in almost all branches of Science.
Advanced Regression Models with SAS and R exposes the reader to the modern world of regression analysis. The material covered by this book consists of regression models that go beyond linear regression, including models for right-skewed, categorical and hierarchical observations. The book presents the theory as well as fully worked-out numerical examples with complete SAS and R codes for each regression. The emphasis is on model accuracy and the interpretation of results. For each regression, the fitted model is presented along with interpretation of estimated regression coefficients and prediction of response for given values of predictors. Features: Presents the theoretical framework for each regression. Discusses data that are categorical, count, proportions, right-skewed, longitudinal and hierarchical. Uses examples based on real-life consulting projects. Provides complete SAS and R codes for each example. Includes several exercises for every regression. Advanced Regression Models with SAS and R is designed as a text for an upper division undergraduate or a graduate course in regression analysis. Prior exposure to the two software packages is desired but not required. The Author: Olga Korosteleva is a Professor of Statistics at California State University, Long Beach. She teaches a large variety of statistical courses to undergraduate and master's students. She has published three statistical textbooks. For a number of years, she has held the position of faculty director of the statistical consulting group. Her research interests lie mostly in applications of statistical methodology through collaboration with her clients in health sciences, nursing, kinesiology, and other fields.
I teach management at the university level, and one of the most important topics covered in the curriculum is human behavior in organizations. As an instructor, I have always been fascinated by the behavior of university students. After all, the university is an organization, and students are the largest group in it. One particular behavior that interested me was student withdrawal from courses throughout the semester. This case illustrates how I used count models to study group differences in course withdrawal. As the total number of courses from which a student withdraws during his or her time in university must be a whole number, the analysis was performed using count models. The case study will explain why such models were used and how the analysis was performed.
Functional Form and Heterogeneity in Models for Count Data
Regression Models for Categorical, Count, and Related Variables
Statistical Modeling of Count Data with Over-dispersion Or Zero-inflation Problems
Joint Models for Longitudinal and Time-to-Event Data
Applied Econometrics with R
Negative Binomial Regression

In this study, we will analyze a supply retailing company's data to model the relationship between their customer's past purchase behavior to predict their future online purchase behavior. The data was divided into time periods from 2016: P1-P6(January 31st to July 30th) and P7(July 31st to August 27th). Based on customer's past purchase information from the P1-P6 period, such as money spent, number of cart additions, transactions type, number of unique purchase dates, number of unique purchase skus, number of page views, number browse dates, company size, and number of products purchased, we aim to find if these information could predict the customer's purchase behavior in the P7 period, which is the number of responses the customer responded to emails sent to them during P7. With the response variable as count data, we model the data in R with the Poisson distribution regression with an offset variable. We also model the number of responses out of the number of emails sent using a logistic regression model. For the Poisson model, since there are zero inflation or over-dispersion issues in the response, hurdle model, zero-inflated-poisson (ZIP) model and zero-inflated-negative-binomial (ZINB) model would be used to handle these issues. Model comparisons among the Poisson model with an offset, logistic regression model, hurdle model, ZIP, ZINB is conducted to select the best model to fit the data using the AIC criterion and the cross-validation criterion.
This book presents statistical methods for the analysis of events. The primary focus is on single equation cross section models. The book addresses both the methodology and the practice of the subject and it provides both a synthesis of a diverse body of literature that hitherto was available largely in pieces, as well as a contribution to the progress of the methodology, establishing several new results and introducing new models. Starting from the standard Poisson regression model as a benchmark, the causes, symptoms and consequences of misspecification are worked out. Both parametric and semi-parametric alternatives are discussed. While semi-parametric models allow for robust interference, parametric models can identify features of the underlying data generation process.
In a conversational tone, Regression & Linear Modeling provides conceptual, user-friendly coverage of the generalized linear model (GLM). Readers will become familiar with applications of ordinary least squares (OLS) regression, binary and multinomial logistic regression, ordinal regression, Poisson regression, and loglinear models. The author returns to certain themes throughout the text, such as testing assumptions, examining data quality, and, where appropriate, nonlinear and non-additive effects modeled within different types of linear models. Available with Perusall—an eBook that makes it easier to prepare for class Perusall is an award-winning eBook platform featuring social annotation tools that allow students and instructors to collaboratively mark up and discuss their SAGE textbook. Backed by research and supported by technological innovations developed at Harvard University, this process of learning through collaborative annotation keeps your students engaged and makes teaching easier and more effective. Learn more.
Practical Guide to Logistic Regression covers the key points of the basic logistic regression model and illustrates how to use it properly to model a binary response variable. This powerful methodology can be used to analyze data from various fields, including medical and health outcomes research, business analytics and data science, ecology, fisheries, astronomy, transportation, insurance, economics, recreation, and sports. By harnessing the capabilities of the logistic model, analysts can better understand their data, make appropriate predictions and classifications, and determine the odds of one value of a predictor compared to another. Drawing on his many years of teaching logistic regression, using logistic-based models in research, and writing about the subject, Professor Hilbe focuses on the most important features of the logistic model. Serving as a guide between the author and readers, the book explains how to construct a logistic model, interpret coefficients and odds ratios, predict probabilities and their standard errors based on the model, and evaluate the model as to its fit. Using a variety of real data examples, mostly from health outcomes, the author offers a basic step-by-step guide to developing and interpreting observation and grouped logistic models as well as penalized and exact logistic regression. He also gives a step-by-step guide to modeling Bayesian logistic regression. R statistical software is used throughout the book to display the statistical models while SAS and Stata codes for all examples are included at the end of each chapter. The example code can be adapted to readers' own analyses. All the code is available on the author's website.
Statistical Models Based on Counting Processes
Applied Categorical and Count Data Analysis
Regression & Linear Modeling
Regression Modeling with Actuarial and Financial Applications
An Applied Approach
Discrete Data Analysis with R

Social science and behavioral science students and researchers are often confronted with data that are categorical, count a phenomenon, or have been collected over time. Sociologists examining the likelihood of interracial marriage, political scientists studying voting behavior, criminologists counting the number of offenses people commit, health scientists studying the number of suicides across neighborhoods, and psychologists modeling mental health treatment success are all interested in outcomes that are not continuous. Instead, they must measure and analyze these events and phenomena in a discrete manner. This book provides an introduction and overview of several statistical models designed for these types of outcomes—all presented with the assumption that the reader has only a good working knowledge of elementary algebra and has taken introductory statistics and linear regression analysis. Numerous examples from the social sciences demonstrate the practical applications of these models. The chapters address logistic and probit models, including those designed for ordinal and nominal variables, regular and zero-inflated Poisson and negative binomial models, event history models, models for longitudinal data, multilevel models, and data reduction techniques such as principal components and factor analysis. Each chapter discusses how to utilize the models and test their assumptions with the statistical software Stata, and also includes exercise sets so readers can practice using these techniques. Appendices show how to estimate the models in SAS, SPSS, and R; provide a review of regression assumptions using simulations; and discuss missing data. A companion website includes downloadable versions of all the data sets used in the book.
Functional Form and Heterogeneity in Models for Count Data surveys practical extensions of the Poisson and negative binomial (NB) models that practitioners can employ to refine the specifications or broaden their reach into new situations. The author resolves some inconsistencies of the panel data models with other more familiar results for the linear regression model. Functional Form and Heterogeneity in Models for Count Data is focused on two large issues: the accommodation of overdispersion and heterogeneity in the basic count framework and the functional form of the conditional mean and the extension of models of heterogeneity to models for panel data and sources of correlation across outcomes. The first is more straightforward since, in principle, these are elements of the conditional variance of the distribution of counts that can be analyzed apart from the conditional mean. Robust inference methods for basic models can be relied upon to preserve the validity of estimation and inference procedures. The second feature motivates the development of more intricate models such as the two part, panel and bivariate models presented in the text.
"This entry-level text offers clear and concise guidelines on how to select, construct, interpret, and evaluate count data. Written for researchers with little or no background in advanced statistics, the book presents treatments of all major models using numerous tables, insets, and detailed modeling suggestions. It begins by demonstrating the fundamentals of linear regression and works up to an analysis of the Poisson and negative binomial models, and to the problem of overdispersion. Examples in Stata, R, and SAS code enable readers to adapt models for their own purposes, making the text an ideal resource for researchers working in public health, ecology, econometrics, transportation, and other related fields"--
This book provides the most comprehensive and up-to-date account of regression methods to explain the frequency of events.
Foundations of Linear and Generalized Linear Models
The Oxford Handbook of Panel Data
Statistical Regression Modeling with R
Modern Statistics for Modern Biology
Generalized Linear Models and Extensions, Second Edition
Modeling Count Data

R is a language and environment for data analysis and graphics. It may be considered an implementation of S, an award-winning language initially - veloped at Bell Laboratories since the late 1970s. The R project was initiated by Robert Gentleman and Ross Ihaka at the University of Auckland, New Zealand, in the early 1990s, and has been developed by an international team since mid-1997. Historically, econometricians have favored other computing environments, some of which have fallen by the wayside, and also a variety of packages with canned routines. We believe that R has great potential in econometrics, both for research and for teaching. There are at least three reasons for this: (1) R is mostly platform independent and runs on Microsoft Windows, the Mac family of operating systems, and various ?avors of Unix/Linux, and also on some more exotic platforms. (2) R is free software that can be downloaded and installed at no cost from a family of mirror sites around the globe, the Comprehensive R Archive Network (CRAN); hence students can easily install it on their own machines. (3) R is open-source software, so that the full source code is available and can be inspected to understand what it really does, learn from it, and modify and extend it. We also like to think that platform independence and the open-source philosophy make R an ideal environment for reproducible econometric research.
A far-reaching course in practical advanced statistics for biologists using R/Bioconductor, data exploration, and simulation.
This book provides a concise point of reference for the most commonly used regression methods. It begins with linear and nonlinear regression for normally distributed data, logistic regression for binomially distributed data, and Poisson regression and negative-binomial regression for count data. It then progresses to these regression models that work with longitudinal and multi-level data structures. The volume is designed to guide the transition from classical to more advanced regression modeling, as well as to contribute to the rapid development of statistics and data science. With data and computing programs available to facilitate readers' learning experience, Statistical Regression Modeling promotes the applications of R in linear, nonlinear, longitudinal and multi-level regression. All included datasets, as well as the associated R program in packages nlme and lme4 for multi-level regression, are detailed in Appendix A. This book will be valuable in graduate courses on applied regression, as well as for practitioners and researchers in the fields of data science, statistical analytics, public health, and related fields.
Panel data econometrics has evolved rapidly over the past three decades. The field is of both theoretical and practical importance, and methods to deal with micro- and macroeconomic panel data are in high demand from practitioners. Applications in finance, development, trade, marketing, health, labor, and consumer economics attest to the usefulness of these methods in applied economics. THis book is a comprehensive source on panel data. It contains 20 chapters edited by Professor Badi Baltagi--one of the leading econometricians in the area of panel data econometrics--and authored by renowned experts in the field. The chapters are divided into two sections. Part I examines new developments in theory. It includes panel cointegration, dynamic panel data models, incidental parameters and dynamic panel modeling, and panel data models for discrete choice. The chapters in Part II target applications of panel data, including health, labor, marketing, trade, productivity and macro applications in panels.
Regression Analysis of Count Data
Applied Generalized Linear Models And Multilevel Models in R
How to Navigate Clueless Colleagues, Lunch-Stealing Bosses, and the Rest of Your Life at Work
Ask a Manager
Econometric Theory and an Application to Labor Mobility
Modeling Count Data Using Stata

The past decades have transformed the world of statistical data analysis, with new methods, new types of data, and new computational tools. The aim of Modern Statistics with R is to introduce you to key parts of the modern statistical toolkit. It teaches you: - Data wrangling - importing, formatting, reshaping, merging, and filtering data in R. - Exploratory data analysis - using visualisation and multivariate techniques to explore datasets. - Statistical inference - modern methods for testing hypotheses and computing confidence intervals. - Predictive modelling - regression models and machine learning methods for prediction, classification, and forecasting. - Simulation - using simulation techniques for sample size computations and evaluations of statistical methods. - Ethics in statistics - ethical issues and good statistical practice. - R programming - writing code that is fast, readable, and free from bugs. Starting from the very basics, Modern Statistics with R helps you learn R by working with R. Topics covered range from plotting data and writing simple R code to using cross-validation for evaluating complex predictive models and using simulation for sample size determination. The book includes more than 200 exercises with fully worked solutions. Some familiarity with basic statistical concepts, such as linear regression, is assumed. No previous programming experience is needed.
In longitudinal studies it is often of interest to investigate how a marker that is repeatedly measured in time is associated with a time to an event of interest, e.g., prostate cancer studies where longitudinal PSA level measurements are collected in conjunction with the time-to-recurrence. Joint Models for Longitudinal and Time-to-Event Data: With Applications in R provides a full treatment of random effects joint models for longitudinal and time-to-event outcomes that can be utilized to analyze such data. The content is primarily explanatory, focusing on applications of joint modeling, but sufficient mathematical details are provided to facilitate understanding of the key features of these models. All illustrations put forward can be implemented in the R programming language via the freely available package JM written by the author. All the R code used in the book is available at: http://jmr.r-forge.r-project.org/
Methods of Statistical Model Estimation examines the most important and popular methods used to estimate parameters for statistical models and provide informative model summary statistics. Designed for R users, the book is also ideal for anyone wanting to better understand the algorithms used for statistical model fitting.The text presents algorith
A hands-on introduction to the principles of Bayesian modeling using WinBUGS Bayesian Modeling Using WinBUGS provides an easily accessible introduction to the use of WinBUGS programming techniques in a variety of Bayesian modeling settings. The author provides an accessible treatment of the topic, offering readers a smooth introduction to the principles of Bayesian modeling with detailed guidance on the practical implementation of key principles. The book begins with a basic introduction to Bayesian inference and the WinBUGS software and goes on to cover key topics, including: Markov Chain Monte Carlo algorithms in Bayesian inference Generalized linear models Bayesian hierarchical models Predictive distribution and model checking Bayesian model and variable evaluation Computational notes and screen captures illustrate the use of both WinBUGS as well as R software to apply the discussed techniques. Exercises at the end of each chapter allow readers to test their understanding of the presented concepts and all data sets and code are available on the book's related Web site. Requiring only a working knowledge of probability theory and statistics, Bayesian Modeling Using WinBUGS serves as an excellent book for courses on Bayesian statistics at the upper-undergraduate and graduate levels. It is also a valuable reference for researchers and practitioners in the fields of statistics, actuarial science, medicine, and the social sciences who use WinBUGS in their everyday work.
Using R, JAGS, Python, and Stan
Regression Models for Categorical Dependent Variables Using Stata, Second Edition
Advanced Regression Models with SAS and R
Practical Guide to Logistic Regression
With Applications in R
Count Data Models

This analysis provides a comprehensive account of models and methods to interpret frequency data.
A valuable overview of the most important ideas and results in statistical modeling Written by a highly-experienced author, Foundations of Linear and Generalized Linear Models is a clear and comprehensive guide to the key concepts and results of linearstatistical models. The book presents a broad, in-depth overview of the most commonly usedstatistical models by discussing the theory underlying the models, R software applications,and examples with crafted models to elucidate key ideas and promote practical modelbuilding. The book begins by illustrating the fundamentals of linear models, such as how the model-fitting projects the data onto a model vector subspace and how orthogonal decompositions of the data yield information about the effects of explanatory variables. Subsequently, the book covers the most popular generalized linear models, which include binomial and multinomial logistic regression for categorical data, and Poisson and negative binomial loglinear models for count data. Focusing on the theoretical underpinnings of these models, Foundations ofLinear and Generalized Linear Models also features: An introduction to quasi-likelihood methods that require weaker distributional assumptions, such as generalized estimating equation methods An overview of linear mixed models and generalized linear mixed models with random effects for clustered correlated data, Bayesian modeling, and extensions to handle problematic cases such as high dimensional problems Numerous examples that use R software for all text data analyses More than 400 exercises for readers to practice and extend the theory, methods, and data analysis A supplementary website with datasets for the examples and exercises An invaluable textbook for upper-undergraduate and graduate-level students in statistics and biostatistics courses, Foundations of Linear and Generalized Linear Models is also an excellent reference for practicing statisticians and biostatisticians, as well as anyone who is interested in learning about the most important statistical models for analyzing data.
Modern survival analysis and more general event history analysis may be effectively handled within the mathematical framework of counting processes. This book presents this theory, which has been the subject of intense research activity over the past 15 years. The exposition of the theory is integrated with careful presentation of many practical examples, drawn almost exclusively from the authors'own experience, with detailed numerical and graphical illustrations. Although Statistical Models Based on Counting Processes may be viewed as a research monograph for mathematical statisticians and biostatisticians, almost all the methods are given in concrete detail for use in practice by other mathematically oriented researchers studying event histories (demographers, econometricians, epidemiologists, actuarial mathematicians, reliability engineers and biologists). Much of the material has so far only been available in the journal literature (if at all), and so a wide variety of researchers will find this an invaluable survey of the subject.
This comprehensive guide to Bayesian methods in astronomy enables hands-on work by supplying complete R, JAGS, Python, and Stan code, to use directly or to adapt. It begins by examining the normal model from both frequentist and Bayesian perspectives and then progresses to a full range of Bayesian generalized linear and mixed or hierarchical models, as well as additional types of models such as ABC and INLA. The book provides code that is largely unavailable elsewhere and includes details on interpreting and evaluating Bayesian models. Initial discussions offer models in synthetic form so that readers can easily adapt them to their own data; later the models are applied to real astronomical data. The consistent focus is on hands-on modeling, analysis of data, and interpretations that address scientific questions. A must-have for astronomers, its concrete approach will also be attractive to researchers in the sciences more generally.
Count Time Series
A Generalized Linear Models Approach
Modern Statistics with R
Regression Models for Categorical and Limited Dependent Variables
Bayesian Models for Astrophysical Data
Modeling Count Data Using Poisson and Negative Binomial Models

After reviewing the linear regression model and introducing maximum likelihood estimation, Long extends the binary logit and probit models, presents multinomial and conditioned logit models and describes models for sample selection bias.
An accessible introduction to the use of regression analysis in the social sciences Regression with Social Data: Modeling Continuous and Limited Response Variables represents the most complete and fully integrated coverage of regression modeling currently available for graduate-level behavioral science students and practitioners. Covering techniques that span the full spectrum of levels of measurement for both continuous and limited response variables, and using examples taken from such disciplines as sociology, psychology, political science, and public health, the author succeeds in demystifying an academically rigorous subject and making it accessible to a wider audience. Content includes coverage of: Logit, probit, scobit, truncated, and censored regressions Multiple regression with ANOVA and ANCOVA models Binary and multinomial response models Poisson, negative binomial, and other regression models for event-count data Survival analysis using multistate, multiepisode, and interval-censored survival models Concepts are reinforced throughout with numerous chapter problems, exercises, and real data sets. Step-by-step solutions plus an appendix of mathematical tutorials make even complex problems accessible to readers with only moderate math skills. The book's logical flow, wide applicability, and uniquely comprehensive coverage make it both an ideal text for a variety of graduate course settings and a useful reference for practicing researchers in the field.
Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R is designed for undergraduate students who have successfully completed a multiple linear regression course, helping them develop an expanded modeling toolkit that includes non-normal responses and correlated structure. Even though there is no mathematical prerequisite, the authors still introduce fairly sophisticated topics such as likelihood theory, zero-inflated Poisson, and parametric bootstrapping in an intuitive and applied manner. The case studies and exercises feature real data and real research questions; thus, most of the data in the textbook comes from collaborative research conducted by the authors and their students, or from student projects. Every chapter features a variety of conceptual exercises, guided exercises, and open-ended exercises using real data. After working through this material, students will develop an expanded toolkit and a greater appreciation for the wider world of data and statistical modeling. A solutions manual for all exercises is available to qualified instructors at the book's website at www.routledge.com, and data sets and Rmd files for all case studies and exercises are available at the authors' GitHub repo (https://github.com/proback/BeyondMLR)
Analysis of Variance, Design, and Regression: Linear Modeling for Unbalanced Data, Second Edition presents linear structures for modeling data with an emphasis on how to incorporate specific ideas (hypotheses) about the structure of the data into a linear model for the data. The book carefully analyzes small data sets by using tools that are easily scaled to big data. The tools also apply to small relevant data sets that are extracted from big data. New to the Second Edition Reorganized to focus on unbalanced data Reworked balanced analyses using methods for unbalanced data Introductions to nonparametric and lasso regression Introductions to general additive and generalized additive models Examination of homologous factors Unbalanced split plot analyses Extensions to generalized linear models R, Minitab®, and SAS code on the author's website The text can be used in a variety of courses, including a yearlong graduate course on regression and ANOVA or a data analysis course for upper-division statistics students and graduate students from other fields. It places a strong emphasis on interpreting the range of computer output encountered when dealing with unbalanced data.
From wrangling and exploring data to inference and predictive modelling
Longitudinal and Multi-level Modeling
Regression With Social Data

Modeling Continuous and Limited Response Variables
Analysis of Variance, Design, and Regression

Probability and Bayesian Modeling is an introduction to probability and Bayesian thinking for undergraduate students with a calculus background. The first part of the book provides a broad view of probability including foundations, conditional probability, discrete and continuous distributions, and joint distributions. Statistical inference is presented completely from a Bayesian perspective. The text introduces inference and prediction for a single proportion and a single mean from Normal sampling. After fundamentals of Markov Chain Monte Carlo algorithms are introduced, Bayesian inference is described for hierarchical and regression models including logistic regression. The book presents several case studies motivated by some historical Bayesian studies and the authors' research. This text reflects modern Bayesian statistical practice. Simulation is introduced in all the probability chapters and extensively used in the Bayesian material to simulate from the posterior and predictive distributions. One chapter describes the basic tenets of Metropolis and Gibbs sampling algorithms; however several chapters introduce the fundamentals of Bayesian inference for conjugate priors to deepen understanding. Strategies for constructing prior distributions are described in situations when one has substantial prior information and for cases where one has weak prior knowledge. One chapter introduces hierarchical Bayesian modeling as a practical way of combining data from different groups. There is an extensive discussion of Bayesian regression models including the construction of informative priors, inference about functions of the parameters of interest, prediction, and model selection. The text uses JAGS (Just Another Gibbs Sampler) as a general-purpose computational method for simulating from posterior distributions for a variety of Bayesian models. An R package ProbBayes is available containing all of the book datasets and special functions for illustrating concepts from the book.
An Applied Treatment of Modern Graphical Methods for Analyzing Categorical DataDiscrete Data Analysis with R: Visualization and Modeling Techniques for Categorical and Count Data presents an applied treatment of modern methods for the analysis of categorical data, both discrete response data and frequency data. It explains how to use graphical meth
The linear regression model is the most commonly used statistical method in the social sciences. This book considers regression models that are appropriate when the dependent variable is censored, truncated, binary, ordinal, nominal, or count. I refer to these variables as categorical and limited dependent variables (hereafter CLDVs). Until recently, the greatest obstacle in using models for CLDVs was the lack of software that was flexible, stable, and easy to use. This limitation no longer applies since these models can be estimated routinely with standard software. Now, the greatest impediment is the complexity of the models and the difficulty in interpreting the results. The difficulties arise because most models for CLDVs are nonlinear.
This book teaches multiple regression and time series and how to use these to analyze real data in risk management and finance.
Beyond Multiple Linear Regression
Methods of Statistical Model Estimation
Visualization and Modeling Techniques for Categorical and Count Data

This monograph deals with econometric models for the analysis of event counts. The interest of econometricians in this class of models has started in the mid-eighties. After more than one decade of intensive research, the litera ture has reached a level of maturity that calls for a systematic and accessible exposition of the main results and methods. Such an exposition is the aim of the book. Count data models have found their way into the curricula of micro-econometric classes and are available on standard computer software. The basic methods have been used in countless applications in fields such as labor economics, health economics, insurance economics, urban economics, and economic demography, to name but a few. Other, more recent, methods are poised to become standard tools soon. While the book is oriented towards the empirical economists and applied econometrician, it should be useful to statisticians and biometricians as well. A first edition of this book was published in 1994 under the title "Count Data Models - Econometric Theory and an Application to Labor Mobility" . While this edition keeps the character and broad organization of this first edition, and its emphasis on combining a summary of the existing literature with several new results and methods, it is substantially revised and enlarged. Many parts have been completely rewritten and several new sections have New sections include: count data models for dependent processes; been added.
Developed from the authors' graduate-level biostatistics course, Applied Categorical and Count Data Analysis explains how to perform the statistical analysis of discrete data, including categorical and count outcomes. The authors describe the basic ideas underlying each concept, model, and approach to give readers a good grasp of the fundamentals of the methodology without using rigorous mathematical arguments. The text covers classic concepts and popular topics, such as contingency tables, logistic models, and Poisson regression models, along with modern areas that include models for zero-modified count outcomes, parametric and semiparametric longitudinal data analysis, reliability analysis, and methods for dealing with missing values. R, SAS, SPSS, and Stata programming codes are provided for all the examples, enabling readers to immediately experiment with the data in the examples and even adapt or extend the codes to fit data from their own studies. Designed for a one-semester course for graduate and senior undergraduate students in biostatistics, this self-contained text is also suitable as a self-learning guide for biomedical and psychosocial researchers. It will help readers analyze data with discrete variables in a wide range of biomedical and psychosocial research fields.
This second edition of Hilbe's Negative Binomial Regression is a substantial enhancement to the popular first edition. The only text devoted entirely to the negative binomial model and its many variations, nearly every model discussed in the literature is addressed. The theoretical and distributional background of each model is discussed, together with examples of their construction, application, interpretation and evaluation. Complete Stata and R codes are provided throughout the text, with additional code (plus SAS), derivations and data provided on the book's website. Written for the practising researcher, the text begins with an examination of risk and rate ratios, and of the estimating algorithms used to model count data. The book then gives an in-depth analysis of Poisson regression and an evaluation of the meaning and nature of overdispersion, followed by a comprehensive analysis of the negative binomial distribution and of its parameterizations into various models for evaluating count data.