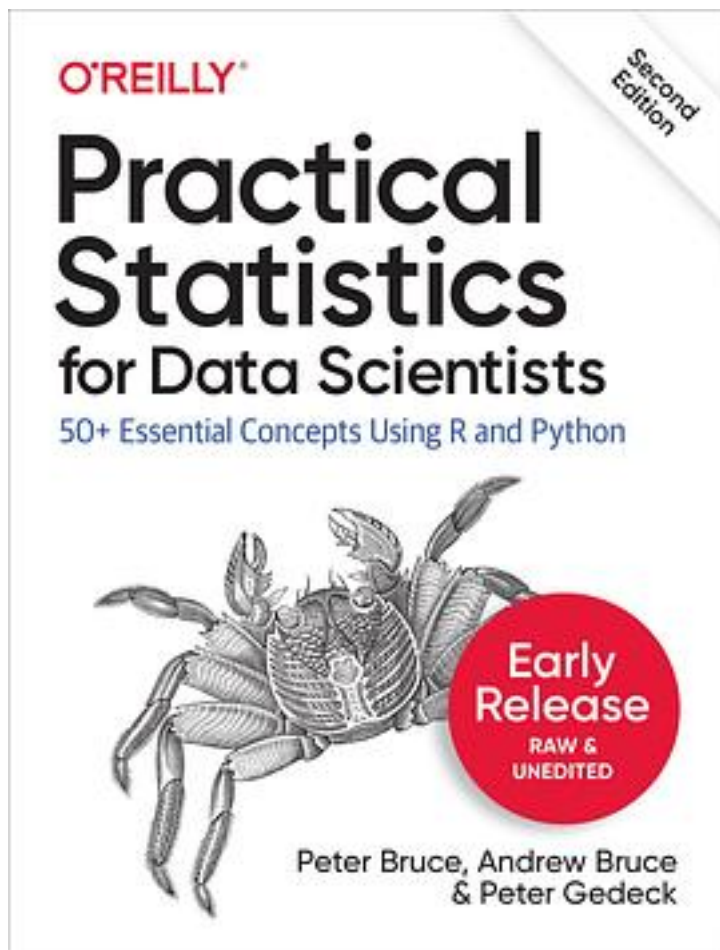


Practical Statistics for Data Scientists, 2nd Edition



[Practical Statistics for Data Scientists, 2nd Edition_下载链接1](#)

著者:Peter Gedeck

出版者:O'Reilly Media, Inc.

出版时间:2020-5-31

装帧:Paperback

isbn:9781492072942

Statistical methods are a key part of data science, yet few data scientists have formal statistical training. Courses and books on basic statistics rarely cover the topic from a data science perspective. The second edition of this practical guide—now including

examples in Python as well as R—explains how to apply various statistical methods to data science, tells you how to avoid their misuse, and gives you advice on what’s important and what’s not.

Many data scientists use statistical methods but lack a deeper statistical perspective. If you’re familiar with the R or Python programming languages, and have had some exposure to statistics but want to learn more, this quick reference bridges the gap in an accessible, readable format.

With this updated edition, you’ll dive into:

Exploratory data analysis

Data and sampling distributions

Statistical experiments and significance testing

Regression and prediction

Classification

Statistical machine learning

Unsupervised learning

作者介绍:

Peter Bruce is the Founder and Chief Academic Officer of the Institute for Statistics Education at Statistics.com, which offers about 80 courses in statistics and analytics, roughly half of which are aimed at data scientists. He has authored or co-authored several books in statistics and analytics, and he earned his Bachelor’s degree at Princeton, and Masters degrees at Harvard and the University of Maryland.

Andrew Bruce, Principal Research Scientist at Amazon, has over 30 years of experience in statistics and data science in academia, government and business. The co-author of Applied Wavelet Analysis with S-PLUS, he earned his bachelor’s degree at Princeton, and PhD in statistics at the University of Washington

Peter Gedeck, Senior Data Scientist at Collaborative Drug Discovery, specializes in the development of machine learning algorithms to predict biological and physicochemical properties of drug candidates. Co-author of Data Mining for Business Analytics, he earned PhD’s in Chemistry from the University of Erlangen-Nürnberg in Germany and Mathematics from Fernuniversität Hagen, Germany

目录:

[Practical Statistics for Data Scientists, 2nd Edition 下载链接1](#)

标签

数据科学

统计实践

统计学

科普

DataScience

数据分析

2020

评论

作为用来准备面试的书，很好。

[Practical Statistics for Data Scientists, 2nd Edition_下载链接1](#)

书评

这本书的作者是统计学领域大咖，Statistics.com统计学教育学院的创立者兼院长，重采样统计软件的开发者。统计学的书市面上有不少了，但能从应用角度把统计学一些关键概念讲明白的不多。虽然书名说是”面向数据科学家“的，但适合所有人用来学习和巩固统计学基础。最好了解一...

真正的问题在于，我们希望p值能包含更多的意义，并且希望p值能够表达如下信息。结果由随机所导致的概率。而且我们希望该值越低越好，这样就可以得出某一假设得到证明的结论。这也是不少期刊编辑对p值的解释。但p值实际所表示的是如下含义。给定一个随机模型，模型所给出的结果...

[Practical Statistics for Data Scientists, 2nd Edition_下载链接1](#)