Estatística Computacional

Computational Statistics

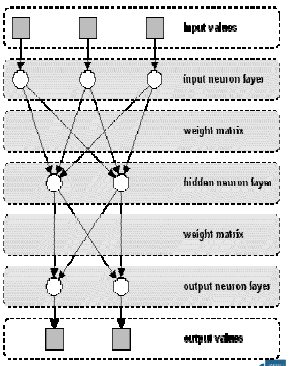
Statistical Computing

Prof. Lori Viali, Dr.
viali@mat.ufrgs.br
http://www.mat.ufrgs.br/~viali/

A tecnologia computacional atual tornou possível grandes avanços na análise de dados, nas várias áreas do conhecimento. Este impacto é mais evidente quando se observa a forma fácil e rápida com que os computadores podem analisar grandes quantidades de dados complexos.

As técnicas multivariadas foram desenvolvidas antes do surgimento dos computadores. Entretanto permaneceram praticamente desconhecidas fora da área da pesquisa. Foi só depois do surgimento dos computadores que elas começaram a ser conhecidas.

Graças aos computadores novas áreas de aplicação e pesquisa se desenvolveram e estão se desenvolvendo como a mineração de dados, as redes neurais, a reamostragem, o reconhecimento de padrões, entre outras.



Exercício

Listar pelo menos cinco áreas da Estatística Computacional

Thomas Bayes
1702 - 1761

*Áreas como a Estatística Bayesiana ganharam novo impulso, graças aos computadores e a **MCMC** (Markow Chain Monte Carlo).*

Métodos de Regularização (smoothing) como o Kernel, Séries Ortogonais, Função Spline, Regressão Não Paramétrica e Regressão Localmente Ponderada.

Métodos de Resampling (estimação) como o Bootstrapping e o Jackknife;

MCMC: Gibbs sampler, algoritmo de Metropolis-Hastings, Simulated Annealing.

A Estatística Computacional é uma interface entre a Estatística, a Computação e a Análise Numérica. O objetivo é desenvolver algoritmos numéricos que sejam eficientes e estáveis, para serem utilizados em cálculos estatísticos, metodologias para bases de dados estatísticas e aplicativos (pacotes).

James E. Gentle (Professor de Estatística Computacional da Universidade George Mason)

Computational Statistics is the area of specialization within statistics that includes statistical visualization and other computationally-intensive methods of statistics. Computational statistics is built on the mathematical theory and methods of statistics, and includes visualization, statistical computing, and Monte Carlo methods. The emphasis in computational statistics is often on exploratory methods.

http://mason.gmu.edu/~jgentle/

Exercício:

Encontrar outras definições para a Estatística Computacional.

Os avanços no hardware e no software alteraram de forma significativa o trabalho das pessoas e em especial a dos Estatísticos. Analistas de Dados e Estatísticos Aplicados dependem dos computadores para armazenar e analisar dados e para produzir relatórios descrevendo as análises feitas.

Estatísticos Matemáticos e Probabilistas utilizam o computador para computação simbólica, para a avaliação de expressões, para a execução de simulações e para a produção de artigos e relatórios.

Alguns efeitos sobre a Estatística ou Estatísticos tem sido sutil como a mudança da utilização de "valores críticos" para os "valores-p", enquanto outros tem sido mais fundamentais como o uso de modelos multivariados ou não lineares sobre o modelo linear univariado, que era utilizado como uma aproximação, uma vez que os multivariados não eram viáveis (James E. Gentle).

Exercício:

Leitura do Artigo: a Estatística está

Associações

ISI (INTERNATIONAL STATISTICAL INSTITUTE)

The International Statistical Institute (ISI) is one of the oldest international scientific associations functioning in the modern world. The Institute is an autonomous society which seeks to develop and improve statistical methods and their application through the promotion of international activity and co-operation.

Multilingual Glossary of Statistical Terms

Johan Dragt of the ISI Senior Executive Corps volunteered as a project leader and has since 1993 compiled a large amount of information in 19 languages for the project. He has constructed a database that contains not only over 3,000 statistical terms, as well as short descriptive texts, including formulae.

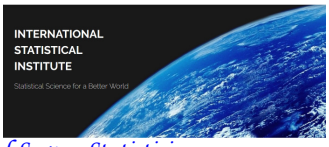
Multilingual Glossary of Statistical Terms

Estatística Computacional

The present version of the multilingual statistical glossary is a first start for Internet usage and is intended to demonstrate the versatility of the information gathered and also provides a means for correcting and expanding the translations provided here with the help of the users of the glossary.

ISI Sections

Estatística Computacional



Bernoulli Society

International Association of Survey Statisticians

International Association for Statistical Computing - IASC

International Association for Official Statistics

International Association for Statistical Education

IASC (International Association for Statistical Computing)

Estatística Computacional

The International Association for Statistical Computing (IASC) was founded during the 41st Session of the ISI in 1977, as a Section of the ISI.

IASC (International Association for Statistical Computing)

Estatística Computacional

The objectives of the Association are to foster world-wide interest in effective statistical computing and to exchange technical knowledge through international contacts and meetings between statisticians, computing professionals, organizations, institutions, governments and the general public.

IASC (International Association for Statistical Computing)

Estatística Computacional

The IASC publishes the Journal Computational Statistics & Data Analysis and organizes its own Conferences (IASC World Conferences, COMPSTAT in Europe, ARS Conferences on Statistical Computing in Asia) and Summer Schools as well as sections of the ISI Conferences.

ASA Section on Statistical Computing and Statistical Graphics

Estatística Computacional

The ASA Section on Statistical Computing was established on January 1, 1972. The mission is to promote the application of computer to statistical problems. We encourage the joint application of statistical techniques and computer technology in other fields, and act as a focal point for computer-oriented activities within the ASA. The importance of these activities is becoming even more critical with the advance of information technologies that challenge us with new problems and enormous data sets.

Statistical Computing Section of the RSS

The section is interested in all aspects of the use of computers and computing in modern statistical analysis. Computational issues are an essential part of modern statistics, data analysis and data mining. Competent statisticians and practitioners must not just understand the principles on which statistical methods work but also be able to implement them on computers in ways that are transparent, require minimal user intervention and scale well with large data sets.

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

日本計算機統計学会
Japanese Society of Computational Statistics

The Japanese Society of Computational Statistics (JSCS) is a scientific and educational society founded in 1986 to broaden and extend excellence of computational statistics, which plays a significant role in the information age, through academic activities for various fields in statistical science and computational science.

The Centre for Computational Statistics and Machine Learning (CSML) spans three departments at University College London, Computer Science, Statistical Science, and the Gatsby Computational Neuroscience Unit.

Conferências

22nd International Conference on COMPUTATIONAL STATISTICS (COMPSTAT 2016)
23-26 August 2016, Auditorio Principe Felipe, Oviedo, Spain

Aims and Scope

The conference aims at bringing together researchers and practitioners to discuss recent developments in computational methods, methodology for data analysis and applications in statistics. All topics within the broad interface of Computing & Statistics will be considered for oral and poster presentation. Topics includes, but are not limited to:

- biostatistics & biocomputing
- categorical data analysis
- clustering & classification
- computer-aided data analysis
- computational Bayesian methods
- computational econometrics
- data visualization
- extreme value theory & applications
- functional data analysis
- high-dimensional data analysis
- kernel & Monte Carlo methods
- machine learning
- mixture models
- multivariate data analysis
- nonparametric statistics
- numerical methods in statistics
- optimization heuristics in statistical modeling
- parametric & semi parametric models
- robust statistics
- sampling methods
- signal processing
- spatial statistics
- symbolic data analysis
- time series analysis

LACSC
I Latin American Conference on Statistical Computing

I LACSC
July 22-24, 2016 - Gramado, State of Rio Grande do Sul, Brazil

Pacific Conference on Statistical Computing and Data Mining

Palm Springs, California - May 6 - May 8

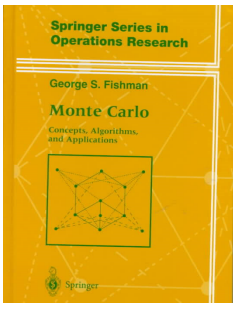
Big Data | Statistics | Algorithms | Resampling Methods | Regression Trees | Computer Science | Programming | Open-source

CCSDA 2016 : 18th International Conference on Computational Statistics and Data Analysis
Paris, France - July 25 - 26, 2016

The ICCSDA 2016: 18th International Conference on Computational Statistics and Data Analysis aims to bring together leading academic scientists, researchers and research scholars to exchange and share their experiences and research results about all aspects of Computational Statistics and Data Analysis. It also provides the premier interdisciplinary forum for researchers, practitioners and educators to present and discuss the most recent innovations, trends, and concerns, practical challenges encountered and the solutions adopted in the field of Computational Statistics and Data Analysis.

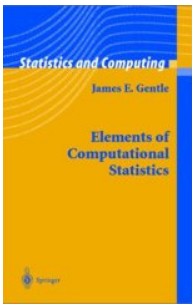
- Advances in Mixture Models
- Analysis of Symbolic and Structured Data
- Applications in Macro-Economics, Finance and Marketing
- Association Rules
- Bayesian Computation methods
- Biostatistics and Bio-computing
- Business Intelligence
- Categorical Data Analysis
- Classification and Discrimination
- Clinical Trials
- Clustering
- Computational Bayesian Methods
- Computational Econometrics
- Computational Methods for Industry
- Computational Methods in Official Statistics
- Computational Statistics Teaching
- Computer-intensive methods for dependent data
- Data Analysis Techniques for Evaluating Trading Strategies
- Data assimilation and its application
- Data Mining
- Data Streams and Massive data
- Design of Experiments
- Dimensionality Reduction
- Econometrics and Statistical Finance
- Environmental Statistics and Climate Change
- Financial econometrics
- Flexible function estimation in high dimensional problems
- Functional Data Analysis
- Functional Genomics: Computational and Statistical Aspects

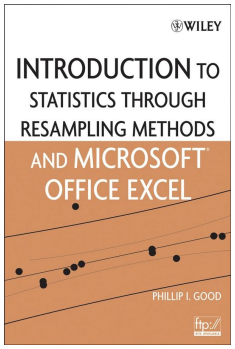
LIVROS



FISHMAN, George. S. Monte Carlo: Concepts, Algorithms and Applications. New York (NY): Springer, 1996. 698 p.

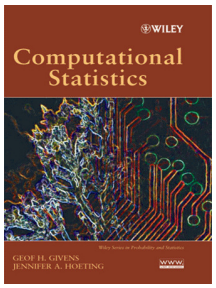
GENTLE, James E. **Elements of Computational Statistics.** London: Springer, 2002. 440 p.

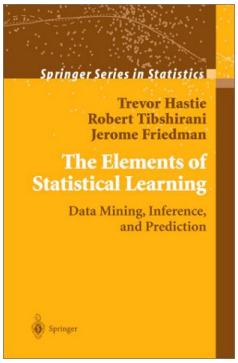




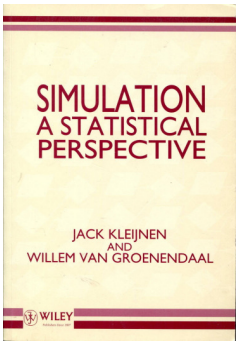
GOOD, Phillip I. Introduction to Statistics through resampling methods and Microsoft Excel. Hoboken (NJ): Willey, 2005. 231 p.

GIVENS, Geof H., HOETING, Jennifer A. **Computational Statistics.** New York (NY): Wiley-Interscience, 2005. 418 p.

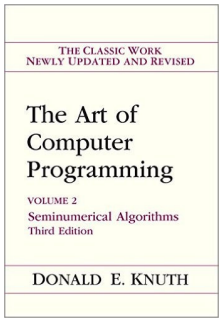




HASTIE, Trevor, TIBSHIRANI, Robert, FRIEDMAN, Jerome. The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York (NY): Springer, Second Edition, 2009 (Disponível na Rede Free).

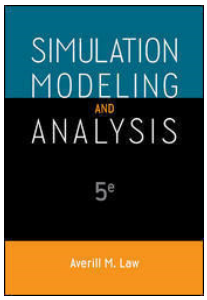


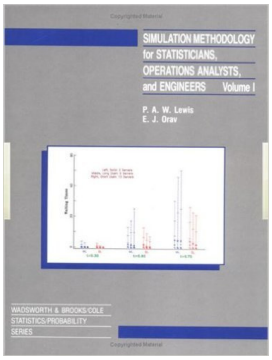
KLEIJNEN, Jack P. C., GROENENDAAL, Willem Van Simulation: a statistical perspective. an. West Sussex (England): John Willey & Sons, 1992. (Disponível no Research Gate)



KNUT, D. The Art of Computer Programming (vol. 2). Reading (MA): Addison-Wesley, 1981.

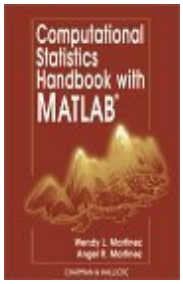
LAW, Averill, KELTON, W. David. Simulation Modeling and Analysis. New York (NY): McGraw-Hill, 2000, Fifth Edition.

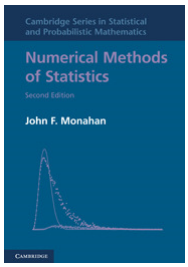




LEWIS, P. A. W., ORAV. E. J. Simulation Methodology for Statisticians, Operation Analysts and Engineers. Belmont (CA): Wadsworth, 1989.

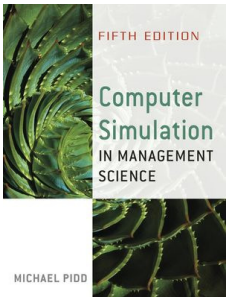
MARTINEZ, Wendy L., MARTINEZ, Angel R. **Computational Statistics Handbook with MATLAB.** New York (NY): Chapman & Hall/CRC. 2001, 616 p.

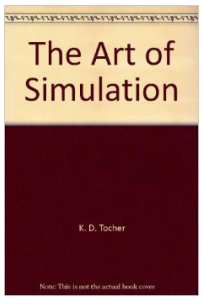




MONAHAN, John F. **Numerical Methods of Statistics.** Cambridge (MA): Cambridge University Press, 2011. 465 p. Second Edition.

PIDD, M. Computer Simulation in Management Science. New York: John Wiley & Sons, April 2006, 5th Edition, 307 p.





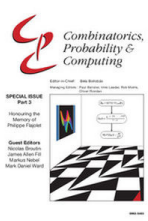
TOCHER, K. D. The Art of Simulation. London: Von Nostrand, 1963. 184 p.

Periódicos (Journals)

Combinatorics, Probability & Computing

Communications in Statistics – Simulation and Computation

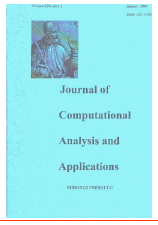
Computacional Statistics



Computational Statistics & Data Analysis

Journal of Computational Analysis and Applications

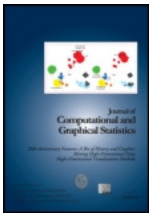
Journal of Informetrics

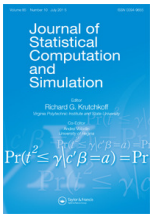


Journal of Computational and Graphical Statistics

Journal of Statistical Computation and Simulation

Journal of Statistical Software





SIAM Journal on Scientific Computing

Statistics and Computing

Pessoas

James Gentle
University Professor of Computational Statistics

GEORGE MASON UNIVERSITY

http://mason.gmu.edu/~jgentle/

Livros de James Gentle

Dr. Hossein Arsham

University of Baltimore

Baltimore, Maryland, USA

http://home.ubalt.edu/ntsbarsh/

Simulações

Let's Make a Deal

ou

The Monty Hall Problem

ou

The Three Doors Puzzle

Enunciado

http://www.prof2000.pt/users/esmmat/portas.htm

Conexões

http://math.rice.edu/~pcmi/mathlinks/montyurl.html

Solução

http://marilynvossavant.com/game-show-problem/

Simulação

3D Normal Distribution in Virtual Reality

The left mouse button changes the point of view

The right mouse button stops and starts the rotation

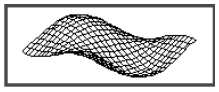
The slider modifies the Correlation Coefficient "r"

http://www.stat.unipg.it/iasc/JAVA/nor3d-en.html

Sites

Center for Computational Statistics

http://www.galaxy.gmu.edu/stats/center.html



Statistics, Statistical Computing, and Mathematics

http://www.pibburns.com/statmath.htm

Computational Statistics in the Data Sciences Program

http://www.scs.gmu.edu/~jgentle/compstat/index.html

The R Project for Statistical Computing

https://www.r-project.org/