

# Index

- $\chi^2$  distribution, [30](#)
- accept-reject sampling, 421
- accessible, [387](#)
- actions, 193
- acyclic, 266
- additive regression, [323](#)
- adjacent, 281
- adjusted treatment effect, 259
- admissibility
  - Bayes rules, 202
- admissible, [202](#)
- AIC (Akaike Information Criterion),  
220
- Aliens, 271
- alternative hypothesis, 95, 149
- ancestor, 265
- aperiodic, 390
- arcs, 281
- associated, 239
- association, 253
  - association is not causation, *16.1*,
- assume, 8
- asymptotic Normality, [128](#)
- asymptotic theory, 71
- asymptotically Normal, [92](#), [126](#)
- asymptotically optimal, [126](#)
- asymptotically uniformly integrable,  
81
- average causal effect, 252
- average treatment effect, 252
- Axiom 1, 5
- Axiom 2, 5
- Axiom 3, 5
- axioms of probability, 5
- backfitting, [324](#)
- bagging, [375](#)
- bandwidth, [313](#)
- Bayes classification rule, [351](#)
- Bayes Estimators, 197
- Bayes risk, [195](#)

- Bayes rules, [197](#)
  - admissibility, [202](#)
- Bayes' Theorem, [12](#), [1.17](#), [12](#)
- Bayesian inference, [89](#), [175](#)
  - strengths and weaknesses, [185](#)
- Bayesian network, [263](#)
- Bayesian philosophy, [175](#)
- Bayesian testing, [184](#)
- Benjamini and Hochberg, [10.26](#), [167](#)
- Benjamini-Hochberg (BH) method, [167](#)
- Bernoulli distribution, [26](#), [29](#)
- Beta distribution, [30](#)
- bias-variance tradeoff, [305](#)
- Bibliographic Remarks, [13](#)
- Binomial distribution, [26](#)
- bins, [303](#), [306](#)
- binwidth, [306](#)
- bivariate distribution, [31](#)
- Bonferroni method, [166](#)
- boosting, [375](#)
- bootstrap, [107](#)
  - parametric, [134](#)
- Bootstrap Confidence Intervals, [110](#)
- bootstrap percentile interval, [111](#)
- bootstrap pivotal confidence, [111](#)
- Bootstrap variance estimation, [109](#)
- branching process, [398](#)
- candidate, [411](#)
- Cauchy distribution, [30](#)
- Cauchy-Schwartz inequality, [4.8](#), [66](#)
- causal odds ratio, [252](#)
- causal regression function, [256](#)
- causal relative risk, [253](#)
- Central Limit Theorem (CLT), [5.8](#), [77](#)
- Chapman-Kolmogorov equations, [23.9](#), [385](#)
- Chebyshev's inequality, [4.2](#), [64](#)
- checking assumptions, [135](#)
- child, [265](#)
- classes, [387](#)
- classification, [349](#)
- classification rule, [349](#)
- classification trees, [360](#)
- classifier
  - assessing error rate, [362](#)
- clique, [285](#)
- closed, [388](#)
- CLT, [77](#)
- collider, [265](#)
- comparing risk functions, [194](#)
- complete, [281](#), [328](#)
- composite hypothesis, [151](#)
- Computer Experiment, [16](#), [17](#)
- concave, [66](#)
- conditional causal effect, [255](#)
- conditional distribution, [36](#)
- conditional expectation, [54](#)
- conditional independence, [264](#)
  - minimal, [287](#)
- conditional likelihood, [213](#)
- Conditional Probability, [10](#)
- conditional probability, [10](#), [10](#)
- conditional probability density function, [37](#)
- conditional probability mass function, [36](#)
- conditioning by intervention, [274](#)
- conditioning by observation, [274](#)
- confidence band, [99](#)
- confidence bands, [323](#)
- confidence interval, [65](#), [92](#)
- confidence set, [92](#)
- confounding variables, [257](#)
- conjugate, [179](#)

- consistency relationship, 252
- consistent, [90](#), [126](#)
- continuity of probabilities, *1.8*, 7
- continuous, [23](#)
- converges in distribution, 72
- converges in probability, 72
- convex, 66
- correlation, [52](#)
  - confidence interval, 234
- cosine basis, 329
- counterfactual, 251, 252
- counting process, 395
- covariance, [52](#)
- covariance matrix, 232
- covariate, 209
- coverage, 92
- critical value, 150
- cross-validation, [363](#)
- cross-validation estimator of risk, [310](#)
- cumulative distribution function, [20](#)
- curse of dimensionality, [319](#)
- curve estimation, 89, [303](#)
- d-connected, 270
- d-separated, 270
- DAG, 266
- data mining, vii
- decision rule, 193
- decision theory, 193
- decomposition theorem, *23.17*, 389
- delta method, *5.13*, 79, [131](#)
- density estimation, 312
  - kernel approach, 312
  - orthogonal function approach, 331
- dependent, [34](#), 239
- dependent variable, 89
- derive, 8
- descendant, 265
- detail coefficients, 342
- detailed balance, 391, 413
- deviance, [299](#)
- directed acyclic graph, 266
- directed graph, 264
- directed path, 265
- discrete, [22](#)
- discrete uniform distribution, [26](#)
- discrete wavelet transform (DWT), 344
- discriminant function, [354](#)
- discrimination, [349](#)
- disjoint, 5
- distribution
  - $\chi^2$ , [30](#)
  - Bernoulli, [26](#), [29](#)
  - Beta, [30](#)
  - Binomial, [26](#)
  - Cauchy, [30](#)
  - conditional, 36
  - discrete uniform, [26](#)
  - Gaussian, [28](#)
  - Geometric, [26](#)
  - Multinomial, [39](#)
  - multivariate Normal, [39](#)
  - Normal, [28](#)
  - point mass, [26](#)
  - Poisson, [27](#)
  - t, [30](#)
  - Uniform, [27](#)
- Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, *7.5*, 98
- edges, 281
- efficient, [126](#), [131](#)
- elements, 3
- EM algorithm, [144](#)
- empirical distribution function, [97](#)

- empirical error rate, [351](#)
- empirical probability measure, [367](#)
- empirical risk minimization, [352](#), [365](#)
- Epanechnikov kernel, [312](#)
- equal in distribution, 25
- equivariant, [126](#)
- ergodic, [390](#)
- Events, 3
- events, 3
- evidence, 157
- Exercises, 13
- expectation, [47](#)
  - conditional, 54
- expected value, [47](#)
- exponential families, [140](#)
  
- faithful, 270
- false discovery proportion, 166
- false discovery rate, 166
- FDP, 166
- FDR, 166
- feature, 89, 209
- first moment, [47](#)
- first quartile, 25
- Fisher information, [128](#)
- Fisher information matrix, [133](#)
- Fisher linear discriminant function, [356](#)
- fitted line, 210
- fitted values, 210
- frequentist (or classical), 175
- frequentist inference, 89
  
- Gamma function, 29
- Gaussian classifier, 353
- Gaussian distribution, [28](#)
- Geometric distribution, [26](#)
- Gibbs sampling, 416
- Gini index, [361](#)
  
- Glivenko-Cantelli theorem, 7.4, 98
- goodness-of-fit tests, 168
- graphical, [294](#)
- graphical log-linear models, 294
  
- Haar father wavelet, 340
- Haar scaling function, 340
- Haar wavelet regression, 343
- hierarchical log-linear model, [296](#)
- hierarchical model, 56
- hierarchical models, 416
- histogram, 303, [305](#)
- histogram estimator, [306](#)
- Hoeffding's inequality, 4.4, 64, [365](#)
- homogeneous, 384
- homogeneous Poisson process, [396](#)
- Horwitz-Thompson, 188
- hypothesis testing, 94
  
- identifiable, [126](#)
- importance sampling, 408
- impurity, [360](#)
- inadmissible, [202](#)
- independent, 8, [8](#), [34](#)
- Independent Events, 8
- independent random variables, 34
- independent variable, 89
- index set, 381
- indicator function, 5
- inequalities, 63
- inner product, 327
- integrated squared error (ISE), [304](#)
- intensity function, [395](#)
- interarrival times, 396
- intervene, 273
- intervention, 273
- Introduction, 3
- invariant, [390](#)
- inverse Gaussian distribution, 421

- irreducible, 388
- iterated expectations, *3.24*, 55
- jackknife, 115
- James-Stein estimator, 204
- Jeffreys-Lindley paradox, 192
- Jensen's inequality, *4.9*, 66
- joint mass function, 31
- K-fold cross-validation, 364
- k-nearest-neighbors, 375
- kernel, 312
- kernel density estimator, 312, 313
- kernelization, 371
- Kolmogorov-Smirnov test, 245
- Kullback-Leibler distance, 126
- Laplace transform, 56
- large sample theory, 71
- law of large numbers, 72
- law of total probability, *1.16*, 12
- lazy, *3.6*, 48
- least favorable prior, 198
- least squares estimates, 211
- leave-one-out cross-validation, 220
- leaves, 361
- Legendre polynomials, 329
- length, 327
- level, 150
- likelihood function, 122
- likelihood ratio statistic, 164
- likelihood ratio test, 164
- limit theory, 71
- limiting distribution, 391
- linear algebra notation, 231
- linear classifier, 353
- linearly separable, 369
- log odds ratio, 240
- log-likelihood function, 122
- log-linear expansion, 292
- log-linear model, 286
- log-linear models, 291
- logistic regression, 223
- loss function, 193
- machine learning, vii
- Manalahobis distance, 353
- marginal Distribution, 33
- marginal distribution, 197
- Markov chain, 383, 383
- Markov condition, 267
- Markov equivalent, 271
- Markov's inequality, *4.1*, 63
- maximal clique, 285
- maximum likelihood, 122
- maximum likelihood estimates
  - computing, 142
- maximum likelihood estimator
  - consistent, 126
- maximum risk, 195
- mean, 47
- mean integrated squared error (MISE), 304
- mean recurrence time, 390
- mean squared error, 91
- measurable, 13, 43
- median, 25
  - bootstrap, 109
- Mercer's theorem, 373
- method of moments estimator, 121
- Metropolis within Gibbs, 419
- Metropolis-Hastings algorithm, 411
- Mill's inequality, *4.7*, 65
- minimal conditional independence, 287
- minimal sufficient, 138
- minimax rule, 197, 198
- missing data, 187

- mixture of Normals, 143
- model generator, 297
- model selection, 218
- moment generating function, [56](#)
- moments, 49
- monotone decreasing, 5
- monotone increasing, 5
- Monte Carlo integration, 404
- Monte Carlo integration method, 404
- Monty Hall, 14
- most powerful, 152
- mother Haar wavelet, 341
- MSE, 91
- Multinomial, 235
- Multinomial distribution, [39](#)
- multiparameter models, [133](#)
- multiple regression, 216
- multiple testing, 165
- multiresolution analysis, 341
- Multivariate central limit theorem, [5.12](#), 78
- Multivariate Delta Method, [5.15](#), 79
- multivariate Normal, 234
- multivariate Normal distribution, [39](#)
- mutually exclusive, 5
- Nadaraya-Watson kernel estimator, [319](#)
- naive Bayes classifier, [359](#)
- natural parameter, [141](#)
- natural sufficient statistic, [140](#)
- neural networks, [376](#)
- Newton-Raphson, 143
- Neyman-Pearson, [10.30](#), 170
- nodes, 281
- non-collider, 265
- non-null, 390
- nonparametric model, 88
- nonparametric regression, [319](#)
  - kernel approach, 319
  - orthogonal function approach, 337
- norm, 327
- normal, 327
- Normal distribution, [28](#)
- Normal-based confidence interval, [6.16](#), 94
- normalizing constant, 177, 403
- not, 10
- nuisance parameter, [120](#)
- nuisance parameters, 88
- null, 390
- null hypothesis, 94, 149
- observational studies, 257
- odds ratio, [240](#)
- olive statistics, i
- one-parameter exponential family, [140](#)
- one-sided test, 151
- optimality, [130](#)
- orthogonal, 327
- orthogonal functions, 327
- orthonormal, 328
- orthonormal basis, 328
- outcome, 89
- overfitting, 218
- p-value, 156, 157
- pairwise Markov graph, 283
- parameter of interest, [120](#)
- parameter space, 88
- parameters, 26
- parametric bootstrap, [134](#)
- parametric model, 87
- parent, 265

- Parseval's relation, 329
- partition, 5
- path, 281
- Pearson's  $\chi^2$  test, 241
- period, 390
- periodic, 390
- permutation distribution, 162
- permutation test, 161
- permutation test:algorithm, 163
- perpendicular, 327
- persistent, 388
- pivot, 110
- plug-in estimator, 99
- point estimation, 90
- point mass distribution, 26
- pointwise asymptotic, 95
- Poisson distribution, 27
- Poisson process, 394, 395
- positive definite, 231
- posterior, 176
  - large sample properties, 181
- posterior risk, 197
- potential, 285
- potential outcomes, 251
- power function, 150
- precision matrix, 232
- predicted values, 210
- prediction, 89, 215
- prediction interval, *13.11*, 215
- prediction risk, 219
- predictor, 89
- predictor variable, 209
- prior distribution, 176
- Probability, 5
- probability, 5
- probability distribution, 5, 5
- probability function, 22
- probability inequalities, 63
- probability mass function, 22
- probability measure, 5, 5
- Probability on Finite Sample Spaces, 7
- proposal, 411
- quadratic discriminant analysis (QDA), 353
- quantile function, 25
- quantiles, 102
- random variable, 19
  - independent, 34
- random vector, 38, 232
- random walk, 59
- random-walk-Metropolis-Hastings, 415
- realizations, 3
- recurrence time, 390
- recurrent, 388
- regression, 89, 209, 335
  - nonparametric, 319
- regression function, 89, 209, 351
- regression through the origin, 226
- regressor, 89
- rejection region, 150
- relative risk, 248
- represents, 266
- residual sums of squares, 211
- residuals, 210
- response variable, 89, 209
- reweighted least squares, 224
- risk, 194, 304
- rule of the lazy statistician, *3.6*, 48
- Rules of d-separation, 270
- sample correlation, 102
- sample mean, 51
- sample outcomes, 3

- sample quantile, 102
- sample space, 3
- Sample Spaces and Events, 3
- sample variance, 51
- sampling distribution, 90
- saturated model, 298, 299
- scaling coefficient, 342
- score function, [128](#)
- se, 90
- shatter coefficient, [367](#)
- shattered, [367](#)
- simple hypothesis, 151
- simple linear regression, 210
- Simpson's paradox, 259
- simulation, 108, 180
- size, [150](#)
- slack variables, [371](#)
- Slutzky's theorem, 75
- smoothing, [303](#)
- smoothing parameter, [303](#)
- Sobolev space, 88
- sojourn times, 396
- spatially inhomogeneous, 340
- standard deviation, [51](#)
- standard error, 90
- standard Normal distribution, 28
- state space, 381
- stationary, [390](#)
- statistic, 61, 107, [137](#)
- statistical functional, 89, 99
- statistical model, 87
- Stein's paradox, 204
- stochastic process, 381
- Stone's theorem, *20.16*, 316
- strong law of large numbers, *5.18*,  
81
- strongly inadmissible, 204
- subjectivism, 181
- sufficiency, 137
- sufficient statistic, [137](#)
- Summary of Terminology, 4
- supervised learning, [349](#)
- support vector machines, [368](#)
- support vectors, [370](#)
- t distribution, [30](#)
- t-test, 170
- test statistic, 150
- third quartile, 25
- thresholding, 342
- training error, 219
- training error rate, [351](#)
- training set, [363](#)
- transformations of random variables,  
41
- transient, [388](#)
- true error rate, [351](#)
- two-sided test, 151
- type I error, 150
- type II error, 150
- types of convergence, 72
- unbiased, 90
- underfitting, 218
- undirected graph, 281
- uniform asymptotic, 95
- Uniform distribution, [27](#)
- unshielded collider, 266
- validation set, [363](#)
- Vapnik-Chervonenkis, [366](#)
- variance, [51](#)  
conditional, [55](#)
- variance-covariance matrix, 53
- vertices, 281
- waiting times, 396
- Wald test, [153](#)



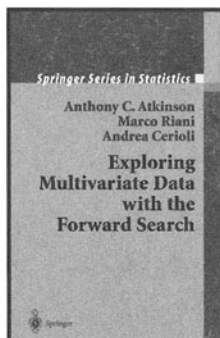
wavelets, 340

weak law of large numbers (WLLN),  
5.6, 76

Zheng-Loh method, 222

- Lehmann*: Testing Statistical Hypotheses, Second Edition  
*Lehmann and Casella*: Theory of Point Estimation, Second Edition  
*Lindman*: Analysis of Variance in Experimental Design  
*Lindsey*: Applying Generalized Linear Models  
*Madansky*: Prescriptions for Working Statisticians  
*McPherson*: Applying and Interpreting Statistics: A Comprehensive Guide, Second Edition  
*Mueller*: Basic Principles of Structural Equation Modeling: An Introduction to LISREL and EQS  
*Nguyen and Rogers*: Fundamentals of Mathematical Statistics: Volume I: Probability for Statistics  
*Nguyen and Rogers*: Fundamentals of Mathematical Statistics: Volume II: Statistical Inference  
*Noether*: Introduction to Statistics: The Nonparametric Way  
*Nolan and Speed*: Stat Labs: Mathematical Statistics Through Applications  
*Peters*: Counting for Something: Statistical Principles and Personalities  
*Pfeiffer*: Probability for Applications  
*Pitman*: Probability  
*Rawlings, Pantula and Dickey*: Applied Regression Analysis  
*Robert*: The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation, Second Edition  
*Robert and Casella*: Monte Carlo Statistical Methods, Second Edition  
*Rose and Smith*: Mathematical Statistics with *Mathematica*  
*Ruppert*: Statistics and Finance: An Introduction  
*Santner and Duffy*: The Statistical Analysis of Discrete Data  
*Saville and Wood*: Statistical Methods: The Geometric Approach  
*Sen and Srivastava*: Regression Analysis: Theory, Methods, and Applications  
*Shao*: Mathematical Statistics, Second Edition  
*Shorack*: Probability for Statisticians  
*Shumway and Stoffer*: Time Series Analysis and Its Applications  
*Simonoff*: Analyzing Categorical Data  
*Terrell*: Mathematical Statistics: A Unified Introduction  
*Timm*: Applied Multivariate Analysis  
*Toutenburg*: Statistical Analysis of Designed Experiments, Second Edition  
*Wasserman*: All of Statistics: A Concise Course in Statistical Inference  
*Whittle*: Probability via Expectation, Fourth Edition  
*Zacks*: Introduction to Reliability Analysis: Probability Models and Statistical Methods

## ALSO AVAILABLE FROM SPRINGER!



### EXPLORING MULTIVARIATE DATA WITH THE FORWARD SEARCH

ANTHONY C. ATKINSON, MARCO RIANI, and ANDREA CERIOLO

This book is about using graphs to explore and model continuous multivariate data. Such data are often modeled using the multivariate normal distribution and there is a literature of weighty statistical tomes presenting the mathematical theory of this activity. This book is very different. It focuses on ways of exploring whether the data do indeed have a normal distribution. Outlier detection, transformations to normality and the detection of clusters and unsuspected influential subsets are emphasized.

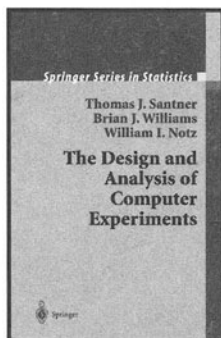
2003/650 PP./HARDCOVER/ISBN 0-387-40852-5  
SPRINGER SERIES IN STATISTICS

### THE DESIGN AND ANALYSIS OF COMPUTER EXPERIMENTS

THOMAS J. SANTNER, BRIAN J. WILLIAMS, and WILLIAM NOTZ

This book describes methods for designing and analyzing experiments conducted using computer code in lieu of a physical experiment. It discusses how to select the values of the factors at which to run the code (the design of the computer experiment) in light of the research objectives of the experimenter. It also provides techniques for analyzing the resulting data so as to achieve these research goals, and it illustrates these methods with code that is available to the reader at the companion web site for the book.

2003/240 PP./HARDCOVER/ISBN 0-387-95420-1  
SPRINGER SERIES IN STATISTICS

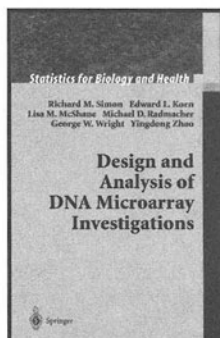


### DESIGN AND ANALYSIS OF DNA MICROARRAY INVESTIGATIONS

RICHARD M. SIMON, EDWARD L. KORN,  
LISA M. MCSHANE, MICHAEL D. RADMACHER,  
GEORGE W. WRIGHT and YINGDONG ZHAO

This book is targeted to biologists with limited statistical background and to statisticians and computer scientists interested in being effective collaborators on multi-disciplinary DNA microarray projects. State-of-the-art analysis methods are presented with minimal mathematical notation and a focus on concepts. This book provides a sound preparation for designing microarray studies that have clear objectives, and for selecting analysis tools and strategies that provide clear and valid answers. The book offers an in depth understanding of the design and analysis of experiments utilizing microarrays and should benefit scientists regardless of what software packages they prefer.

2003/504 PP./HARDCOVER/ISBN 0-387-00135-2  
STATISTICS FOR BIOLOGY AND HEALTH



#### To Order or for Information:

*In the Americas:* **CALL:** 1-800-SPRINGER or  
**FAX:** (201) 348-4505 • **WRITE:** Springer-Verlag New  
York, Inc., Dept. S5634, PO Box 2485, Secaucus, NJ  
07096-2485 • **VISIT:** Your local technical bookstore  
• **E-MAIL:** orders@springer-ny.com

*Outside the Americas:* **CALL:** +49 (0) 6221 345-217/8  
• **FAX:** +49 (0) 6221 345-229 • **WRITE:** Springer  
Customer Service, Haberstrasse 7, 69126  
Heidelberg, Germany • **E-MAIL:** orders@springer.de

PROMOTION: S5634



Springer

www.springer-ny.com