

Bibliography

Literature in Statistics

In the Library of Congress classification scheme, most books on statistics, including statistical computing, are in the QA276 section, although some are classified under H, HA, and HG. Numerical analysis is generally in QA279, and computer science in QA76. Many of the books in the interface of these disciplines are classified in these or other places within QA.

Current Index to Statistics, published annually by the American Statistical Association and the Institute for Mathematical Statistics, contains both author and subject indexes that are useful in finding journal articles or books in statistics. The *Index* is available in hard copy and on CD-ROM.

Mathematical Reviews, published by the American Mathematical Society (AMS), contains brief reviews of articles in all areas of mathematics. The areas of “Statistics”, “Numerical Analysis”, and “Computer Science” contain reviews of articles relevant to computational statistics. The papers reviewed in *Mathematical Reviews* are categorized according to a standard system that has slowly evolved over the years. In this taxonomy, called the AMS MR classification system, “Statistics” is 62Xyy; “Numerical Analysis”, including random number generation, is 65Xyy; and “Computer Science” is 68Xyy. (“X” represents a letter and “yy” represents a two-digit number.) *Mathematical Reviews* is available to subscribers via the World Wide Web at MathSciNet:

<http://www.ams.org/mathscinet/>

Resources Available over the Internet

The best way of storing information is in a digital format that can be accessed by computers. In some cases, the best way for people to access information is by computers; in other cases, the best way is via hard copy, which means that the information stored on the computer must go through a printing process resulting in books, journals, or loose pages.

A huge amount of information and raw data are available online, much in publicly accessible sites. Some of the repositories give space to ongoing

discussions to which anyone can contribute.

For statistics, one of the most useful sites on the Internet is the electronic repository `statlib`, maintained at Carnegie Mellon University, which contains programs, datasets, and other items of interest. The URL is

`http://lib.stat.cmu.edu`

There are two major problems in using the WWW to gather information. One is the sheer quantity of information and the number of sites providing information. The other is the “kiosk problem”; anyone can put up material. Sadly, the average quality is affected by a very large denominator. The kiosk problem may be even worse than a random selection of material; the “fools in public places” syndrome is much in evidence.

It is not clear at this time what will be the media for the scientific literature within a few years. Many of the traditional journals will be converted to an electronic version of some kind. Journals will become Web sites. That is for certain; the details, however, are much less certain. Many bulletin boards and discussion groups have already evolved into electronic journals.

References to the Literature

The following bibliography obviously covers a wide range of topics in statistical computing and computational statistics. Except for a few of the general references, all of these entries have been cited in the text.

The purpose of this bibliography is to help the reader get more information; hence we eschew “personal communications” and references to technical reports that may or may not exist. Those kinds of references are generally for the author rather than for the reader.

In some cases, important original papers have been reprinted in special collections, such as Samuel Kotz and Norman L. Johnson (Editors) (1997), *Breakthroughs in Statistics, Volume III*, Springer-Verlag, New York. In most such cases, because the special collection may be more readily available, we list both sources.

A Note on the Names of Authors

In these references, we have generally used the names of authors as they appear in the original sources. This may mean that the same author will appear with different forms of names, sometimes with given names spelled out, and sometimes abbreviated. In the author index, beginning on page 67, we use a single name for the same author. The name is generally the most unique (i.e., least abbreviated) of any of the names of that author in any of the references. This convention may occasionally result in an entry in the author index that does not occur exactly in any references. For example, a reference to J. Paul Jones together with one to John P. Jones, if we know that the two names refer to the same person, would result in an Author Index entry for John Paul Jones.

- Alam, Khursheed, and Amitava Mitra (1987), Adaptive robust methods for regression analysis based on cross-validation error, *Proceedings of the Business and Economic Statistics Section, ASA*, 643–648.
- Ammann, Larry P. (1989), Robust principal components, *Communications in Statistics — Simulation and Computation* **18**, 857–874.
- Ammann, Larry P. (1993), Robust singular value decompositions: A new approach to projection pursuit, *Journal of the American Statistical Association* **88**, 505–514.
- Ammann, Larry, and John Van Ness (1989), Standard and robust orthogonal regression, *Communications in Statistics — Simulation and Computation* **18**, 145–162.
- Andrews, D. F.; P. J. Bickel; F. R. Hampel; P. J. Huber; W. H. Rogers; and J. W. Tukey (1972), *Robust Estimation of Location: Survey and Advances*, Princeton University Press, Princeton.
- Arnold, Barry C.; Robert J. Beaver; Richard A. Groeneveld; and William Q. Meeker (1993), The nontruncated marginal of a truncated bivariate normal distribution, *Psychometrika* **58**, 471–488.
- Arthanari, T. S., and Yadolah Dodge (1981), *Mathematical Programming in Statistics*, John Wiley & Sons, New York.
- Azzalini, A., and A. Capitanio (2003), Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution, *Journal of the Royal Statistical Society, Series B* **65**, ****
- Balakrishnan, N., and N. Kannan (2003), Variance of a Winsorized mean when the sample contains multiple outliers, *Communications in Statistics — Theory and Methods* **32**, 139–149.
- Barrodale, I., and F. D. K. Roberts (1974), Algorithm 478: Solution of an overdetermined system of equations in the l_1 norm, *Communications of the ACM* **17**, 319–320.
- Bassett, Gilbert W. (1991), Equivalent, monotone, 50% breakdown estimators, *The American Statistician* **45**, 135–137.
- Bassett, Gilbert, and Roger Koenker (1978), Asymptotic theory of least absolute value regression, *Journal of the American Statistical Association* **73**, 618–622.
- Beirlant, Jan; Petra Vynckier; and Jozef L. Teugels (1996), Tail index estimation, Pareto quantile plots, and regression diagnostics, *Journal of the American Statistical Association* **91**, 1659–1667.
- Bickel, P. J., and E. L. Lehmann (1976), Descriptive statistics for nonparametric models. III: Dispersion, *Annals of Statistics* **4**, 1139–1158.
- Birkes, David, and Yadolah Dodge (1993), *Alternative Methods of Regression*, John Wiley & Sons, New York.
- Campbell, N. A. (1980), Robust procedures in multivariate analysis I: Robust covariance estimation, *Applied Statistics* **29**, 231–237.
- Capéraà, Philippe, and Louis-Paul Rivest (1995), On the variance of the trimmed mean, *Statistics and Probability Letters* **22**, 79–85.

- Caroni, C. (2000), Outlier detection by robust principal components analysis, *Communications in Statistics — Simulation and Computation* **29**, 139–151.
- Charnes, A.; W. W. Cooper; and R. O. Ferguson (1955), Optimal estimation of executive compensation by linear programming, *Management Science* **1**, 138–150.
- Chatterjee, Samprit, and Martin Mächler (1997), Robust regression: A weighted least squares approach, *Communications in Statistics — Theory and Methods* **26**, 1381–1394.
- Chaudhuri, Probal (1996), On a geometric notion of quantiles for multivariate data *Journal of the American Statistical Association* **91**, 862–872.
- Cleveland, William S. (1979), Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* **74**, 829–836.
- Cleveland, William S. (1981), LOWESS: A program for smoothing scatterplots by robust locally weighted regression, *The American Statistician* **35**, 54.
- Collins, John R. (2000), Robustness comparisons of some classes of location parameter estimators, *Annals of the Institute of Statistical Mathematics* **52**, 351–366.
- Cook, R. D.; D. M. Hawkins; and S. Weisberg (1993), Exact iterative computation of the robust multivariate minimum volume ellipsoid estimator, *Statistics and Probability Letters* **16**, 213–218.
- Cook, R. Dennis, and Sanford Weisberg (1982), *Residuals and Influence in Regression*, Chapman & Hall, London.
- Cuesta-Albertos, J. A.; A. Catrán Gordaliza; and C. Matrán (1997), Trimmed k -means: An attempt to robustify quantizers, *Annals of Statistics* **25**, 553–576.
- David, H. A., and N. Balakrishnan (1996), Product moments of order statistics and the variance of a lightly trimmed mean, *Statistics and Probability Letters* **29**, 85–87.
- Davies Laurie, and Ursula Gather (1993), The identification of multiple outliers (with discussion), *Journal of the American Statistical Association* **88**, 782–801.
- Davies, P. L. (1993), Aspects of robust linear regression, *Annals of Statistics* **21**, 1843–1899.
- Diaconis, Persi, and David Freedman (1984), Asymptotics of graphical projection pursuit, *The Annals of Statistics* **12**, 793–815.
- Dielman, Terry E.; Cynthia Lowry; and Roger Pfaffenberger (1994), A comparison of quantile estimators, *Communications in Statistics — Simulation and Computation* **23**, 355–371.
- Dielman, Terry E., and Elizabeth L. Rose (1997), A note on hypothesis testing in LAV multiple regression: A small sample comparison, *Computational Statistics & Data Analysis* **23**, 381–388.
- Dodge, Yadolah, and Jana Jureckova (2000), *Adaptive Linear Regression*, Springer-Verlag, New York.
- Dongarra, Jack; Tom Rowan; and Reed Wade (1995), Software distribution using Xnetlib, *ACM Transactions on Mathematical Software* **21**, 79–88.

- Dorfman, A. H. (1989), Iterated reweighted least squares revisited: The tension between simulation and practice, *Computer Science and Statistics: Proceedings of the Twenty-first Symposium on the Interface* (edited by Kenneth Berk and Linda Malone), American Statistical Association, Alexandria, Virginia, 280–283.
- Duarte, Antonio Marcos, and Beatriz Vaz de Melo Mendes (1998), Interior point algorithms for LSAD and LMAD estimation, *Computational Statistics* **13**, 233–256.
- Falk, Michael (1997), On MAD and comedians, *Annals of the Institute of Statistical Mathematics* **49**, 615–644.
- Freimer, Marshall; Govind S. Mudholkar; Georgia Kollia; and Thomas C. Lin (1988), A study of the generalized Tukey lambda family, *Communications in Statistics — Theory and Methods* **17**, 3547–3567.
- García-Escudero, Luis Ángel, and Alfonso Gordaliza (1999), Robustness properties of k means and trimmed k means *Journal of the American Statistical Association* **94**, 956–969.
- García-Escudero, Luis A.; Alfonso Gordaliza; and Carlos Matrán (1999), A central limit theorem for multivariate generalized trimmed k -means *The Annals of Statistics* **27**, 1061–1079.
- Gentle, J. E.; W. J. Kennedy; and V. A. Sposito (1976), Properties of the L_1 -estimate space, *Proceedings of the Statistical Computing Section, ASA*, 163–164.
- Gentle, James E.; Subhash C. Narula; and Vince A. Sposito (1992), Testing software for robust regression, *L_1 Statistical Analysis and Related Methods* (edited by Y. Dodge), Elsevier Science Publishers, Amsterdam, 361–369.
- Gilchrist, Warren G. (2000), *Statistical Modelling with Quantile Functions*, Chapman & Hall / CRC, Boca Raton.
- Gonin, René, and Arthur H. Money (1989), *Nonlinear L_p -Norm Estimation*, Marcel Dekker, Inc., New York.
- Green, P. J. (1984), Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives (with discussion), *Journal of the Royal Statistical Society, Series B* **46**, 149–192.
- Griffiths, P., and I. D. Hill (Editors) (1985), *Applied Statistics Algorithms*, Ellis Horwood Limited, Chichester, United Kingdom.
- Ha, Cu D., and Subhash C. Narula (1989), Perturbation analysis for the minimum sum of absolute errors regression, *Communications in Statistics — Simulation and Computation* **18**, 957–970.
- Hampel, Frank R. (1974) The influence curve and its role in robust estimation, *Journal of the American Statistical Association* **69**, 383–393.
- Hampel, Frank R.; Elvezio M. Ronchetti; Peter J. Rousseeuw; and Werner A. Stahel (1986), *Robust Statistics: The Approach Based on Influence Functions*, John Wiley & Sons, New York.
- Harrell, Frank E., and C. E. Davis (1982), A new distribution-free quantile estimator, *Biometrika* **69**, 635–640.

- Hawkins, Douglas M. (1993a), A feasible solution algorithm for minimum volume ellipsoid estimator in multivariate data, *Computational Statistics* **8**, 95–107.
- Hawkins, Douglas M. (1993b), The feasible set algorithm for least median of squares regression, *Computational Statistics & Data Analysis* **16**, 81–101.
- Hawkins, Douglas M., and David Olive (1999), Applications and algorithms for least trimmed sum of absolute deviations regression, *Computational Statistics & Data Analysis* **32**, 119–134.
- Heiberger, R. M., and R. A. Becker (1992), Design of an S function for robust regression using iteratively reweighted least squares, *Journal of Computational and Graphical Statistics* **1**, 181–196.
- Hesterberg, Timothy C., and Barry L. Nelson (1998), Control variates for probability and quantile estimation, *Management Science* **44**, 1295–1312.
- Hettmansperger, Thomas P., and Joseph W. McKean (1998), *Robust Nonparametric Statistical Models*, Kendall's Library of Statistics 5, John Wiley & Sons, New York.
- Hettmansperger, T. P., and S. J. Sheather (1992), A cautionary note on the method of least median squares, *The American Statistician* **46**, 79–83.
- Hoaglin, David C. (1985), Summarizing shape numerically: The g -and- h distributions, *Exploring Data tables, Trends, and Shapes* (edited by David C. Hoaglin, Frederick Mosteller, and John W. Tukey), John Wiley & Sons, New York, 461–513.
- Hogg, Robert V. (1974), Adaptive robust procedures: A partial review and some suggestions for future applications and theory (with discussion), *Journal of the American Statistical Association* **69**, 909–927.
- Jones, M. C., and M. J. Faddy (2003), A skew extension of the t -distribution, with applications, *Journal of the Royal Statistical Society, Series B* **65**, 159–174.
- Jung, Sin-Ho (1996), Quasi-likelihood methods for median regression models, *Journal of the American Statistical Association* **91**, 251–257.
- Kaigh, W. D., and Peter A. Lachenbruch (1982), A generalized quantile estimator, *Communications in Statistics — Theory and Methods* **11**, 2217–2238.
- Karian, Zaven A., and Edward J. Dudewicz (2000), *Fitting Statistical Distributions*, CRC Press, Boca Raton.
- Koenker, Roger, and Stephen Portnoy (1997), The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators (with discussion), *Statistical Science* **12**, 279–299.
- Léger, Christian, and Joseph P. Romano (1990), Bootstrap adaptive estimation: The trimmed-mean example, *Canadian Journal of Statistics* **18**, 297–314.
- Léger, Christian; Dimitris N. Politis; and Joseph P. Romano (1992), Bootstrap technology and applications, *Technometrics* **34**, 378–398.
- Marazzi, A. (1993), *Algorithms, Routines and S Functions for Robust Statistics* Wadsworth & Brooks/Cole, Pacific Grove, California.
- Maritz, J. S., and R. G. Jarrett (1978), A note on estimating the variance of the sample median, *Journal of the American Statistical Association* **73**,

194–196.

- McKean, Joseph W., and Simon J. Sheather (2000), Partial residual plots based on robust fits, *Technometrics* **42**, 249–261.
- Meintanis, S. G., and G. S. Donatos (1997), A comparative study of some robust methods for coefficient-estimation in linear regression, *Computational Statistics & Data Analysis* **23**, 525–540.
- Morgenthaler, Stephan (1992), Least-absolute-deviations fits for generalized linear models, *Biometrika* **79**, 747–754.
- Narula, Subhash C.; Vince A. Sposito; and John F. Wellington (1993), Intervals which leave the minimum sum of absolute errors regression unchanged, *Applied Statistics* **42**, 369–378.
- Narula, S. C., and J. F. Wellington (1979), Linear regression using multiple-criteria, *Multiple Criteria Decision Making: Theory and Application* (edited by G. Fandel and T. Gal), Springer-Verlag, New York, 266–277.
- Narula, Subhash C., and John F. Wellington (1985), Interior analysis for the minimum sum of absolute errors regression, *Technometrics* **27**, 181–188.
- Nolan, John P., (1998a), Multivariate stable distributions: Approximation, estimation, simulation and identification, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications* (edited by Robert J. Adler, Raisa E. Feldman, and Murad S. Taqqu), Birkhäuser, Boston, 509–526.
- Nolan, John P., (1998b), Univariate stable distributions: Parameterizations and software, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications* (edited by Robert J. Adler, Raisa E. Feldman, and Murad S. Taqqu), Birkhäuser, Boston, 527–533.
- Öztürk, Ömer, and Thomas P. Hettmansperger (1998), Simultaneous robust estimation of location and scale parameters: A minimum-distance approach, *Canadian Journal of Statistics* **26**, 217–229 (Corrections, 1999, *ibid.* **27**, 667).
- Rojo, Javier (1996), On tail categorization of probability laws, *Journal of the American Statistical Association* **91**, 378–384.
- Rousseeuw, P. J. (1984), Least median of squares regression, *Journal of the American Statistical Association* **79**, 871–880. (Reprinted in Samuel Kotz and Norman L. Johnson (Editors) (1997), *Breakthroughs in Statistics, Volume III*, Springer-Verlag, New York, 440–461.)
- Rousseeuw, Peter (1985), Multivariate estimation with high breakdown point, *Mathematical Statistics and Applications* (edited by W. Grossmann, G. C. Pflug, I. Vincze, and W. Wertz), Reidel Publishing, Dordrecht, 283–297.
- Rousseeuw, Peter J., and Mia Hubert (1999), Regression depth (with discussion), *Journal of the American Statistical Association* **94**, 388–433.
- Rousseeuw, Peter J., and Annick M. Leroy (1987), *Robust Regression and Outlier Detection*, John Wiley & Sons, New York.
- Rousseeuw, Peter J., and Ida Ruts (1996), Bivariate location depth, *Applied Statistics* **45**, 516–526.
- Rousseeuw, Peter J., and Ida Ruts (1998), Constructing the bivariate Tukey median, *Statistica Sinica* **8**, 827–839.

- Rousseeuw, P. J., and B. van Zomeren (1990), Unmasking multivariate outliers and leverage points, *Journal of the American Statistical Association* **85**, 633–651.
- Ruppert, David, and Raymond J. Carroll (1980), Trimmed least squares estimation in the linear model, *Journal of the American Statistical Association* **75**, 828–838 (Corrections, 1982, *ibid.* **77**, 954).
- Ruts, Ida, and Peter J. Rousseeuw (1996), Computing depth contours of bivariate point clouds, *Computational Statistics and Data Analysis* **23**, 153–168.
- Schlossmacher, E. J. (1973), An iterative technique for absolute deviations curve fitting, *Journal of the American Statistical Association* **68**, 857–859.
- Schönhage, A.; M. Paterson; and N. Pippenger (1976), Finding the median, *Journal of Computers and System Sciences* **13**, 184–199.
- Seaver, Bill; Konstantinos Triantis; and Chip Reeves (1999), The identification of influential subsets in regression using a fuzzy clustering strategy, *Technometrics* **41**, 340–351.
- Sheather, Simon J., (1987), Assessing the accuracy of the sample median: Estimated standard errors versus interpolated confidence intervals, *Statistical Data Analysis Based on the L_1 -Norm and Related Methods*, Elsevier Science Publishers, Amsterdam, 203–215.
- Sielken, R. L., and H. O. Hartley (1973), Two linear programming algorithms for unbiased estimation of linear models, *Journal of the American Statistical Association* **68**, 639–641.
- Simpson, James R., and Douglas C. Montgomery (1998a), The development and evaluation of alternative generalized M -estimation techniques, *Communications in Statistics — Simulation and Computation* **27**, 999–1018.
- Simpson, James R., and Douglas C. Montgomery (1998b), A performance-based assessment of robust regression methods, *Communications in Statistics — Simulation and Computation* **27**, 1031–1049.
- Singleton, R. R. (1940), A method of minimizing the sum of absolute values of deviations, *Annals of Mathematical Statistics* **11**, 301–310.
- Souvaine, D. L., and J. M. Steele (1987), Time- and space-efficient algorithms for least median of squares regression, *Journal of the American Statistical Association* **82**, 794–801.
- Spanier, Jerome, and Keith B. Oldham (1987), *An Atlas of Functions*, Hemisphere Publishing Corporation, Washington (also Springer-Verlag, Berlin).
- Staudte, Robert G., and Simon J. Sheather (1990), *Robust Estimation and Testing*, John Wiley & Sons, New York.
- Street, James O.; Raymond J. Carroll; and David Ruppert (1988), A note on computing robust regression estimates via iteratively reweighted least squares, *The American Statistician* **42**, 152–154.
- Struyf, Anja, and Peter J. Rousseeuw (2000), High-dimensional computation of the deepest location, *Computational Statistics & Data Analysis* **34**, 415–426.
- Thompson, William J. (1997), *Atlas for Computing Mathematical Functions*:

- An Illustrated Guide for Practitioners with Programs in C and Mathematica*, John Wiley & Sons, New York.
- Turkheimer, F.; K. Pettigrew; L. Sokoloff; and K. Schmidt (1999), A minimum variance adaptive technique for parameter estimation and hypothesis testing, *Communications in Statistics — Simulation and Computation* **28**, 931–956.
- Velilla, Santiago (1995), Diagnostics and robust estimation in multivariate data transformations, *Journal of the American Statistical Association* **90**, 945–951.
- Wang, Ferdinand T., and David W. Scott (1994), The L_1 method for robust nonparametric regression, *Journal of the American Statistical Association* **89**, 65–76.
- Wang, Naisyin, and Adrian E. Raftery (2002), Nearest-neighbor variance estimation (NNVE): Robust covariance estimation via nearest-neighbor cleaning (with discussion), *Journal of the American Statistical Association* **97**, 994–1020.
- Weiss, Robert E. (1992), Influence diagnostics with the Gibbs sampler, *Computing Science and Statistics* **24**, 266–270.
- Wilcox, Rand R., (1997), *Introduction to Robust Estimation and Hypothesis Testing*, Academic Press, San Diego.
- Wilcox, Rand R., and Jan Muska (1999), Tests of hypotheses about regression parameters when using a robust estimator, *Communications in Statistics — Theory and Methods* **28**, 2201–2212.
- Woodruff, David L., and David M. Rocke (1993), Heuristic search algorithms for the minimum volume ellipsoid, *Journal of Computational and Graphical Statistics* **2**, 69–95.
- Woodruff, David L., and David M. Rocke (1994), Computable robust estimation of multivariate location and shape in high dimension using compound estimators, *Journal of the American Statistical Association* **89**, 888–896.
- Xu, Chong-Wei, and Wei-Kei Shiue (1993), Parallel algorithms for least median of squares regression, *Computational Statistics & Data Analysis* **16**, 349–362.
- Yohai, Victor J., and Ruben H. Zamar (1988), High breakdown-point estimates of regression by means of the minimization of an efficient scale, *Journal of the American Statistical Association* **83**, 406–413. (See also *ibid.*, 1989, **84**, 636.)
- Wang, Naisyin, and Adrian E. Raftery (2002), Nearest-neighbor variance estimation (NNVE): Robust covariance estimation via nearest-neighbor cleaning (with discussion), *Journal of the American Statistical Association* **97**, 994–1019.
- Zani, Sergio; Marco Riani; and Aldo Corbellini (1998), Robust bivariate box-plots and multiple outlier detection, *Computational Statistics & Data Analysis* **28**, 257–270.