
References

Methodological References

- Agresti, A. 2002. *Categorical Data Analysis*, 2nd edn. John Wiley.
- Aitchison, J. 2003. *The Statistical Analysis of Compositional Data*. Blackburn Press.
- Aldrich, J. 1995. Correlations genuine and spurious in Pearson and Yule. *Statistical Science* 10: 364–76.
- Ambrose, C. and McLachlan, G. J. 2002. Selection bias in gene extraction on the basis of microarray gene-expression data. *PNAS* 99: 6262–6.
- Andersen, B. 1990. *Methodological Errors in Medical Research: An Incomplete Catalogue*. Blackwell Scientific.
- Barnett, V. 2002. *Sample Survey: Principles & Methods*, 2nd edn. Arnold Publishers.
- Bartholemew, D. 2004. *Measuring Intelligence. Facts and Fallacies*. Cambridge University Press.
- Bates, D. 2005. Fitting linear mixed models in R. *R News* 5(1): 27–30.
- Bates, D. M. and DebRoy, S. 2003. Converting a large R package to S4 classes and methods. In K. Hornik, F. Leisch and A. Zeileis (eds), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/>
- Bates, D. M. and Watts, D. G. 1988. *Nonlinear Regression Analysis and Its Applications*. John Wiley.
- Belson, W. A. 1959. Matching and prediction on the principle of biological classification. *Applied Statistics* 8: 65–75.
- Berk, R. A. 2008. *Statistical Learning from a Regression Perspective*. Springer.
- Bickel, P. J., Hammel, E. A. and O'Connell, J. W. 1975. Sex bias in graduate admissions: data from Berkeley. *Science* 187: 398–403.
- Bland, M. and Altman, D. 2005. Do the left-handed die young? *Significance* 2: 166–70.
- Bolker, B. M. 2008. *Ecological Models and Data in R*. Princeton University Press.
- Box, G. E. P. and Cox, D. R. 1964. An analysis of transformations (with discussion). *Journal of the Royal Statistical Society B* 26: 211–52.
- Braun, W. J. and Murdoch, D. J. 2007. *A First Course in Statistical Programming with R*. Cambridge University Press.
- Breiman, L. 2001. Statistical modeling: the two cultures. *Statistical Science* 16: 199–215.
- Brockwell, P. and Davis, R. A. 2002. *Time Series: Theory and Methods*, 2nd edn. Springer.
- Canty, A. J. 2002. Resampling methods in R: the boot package. *R News* 2/3: 2–7.
- Carroll, R. 2004. Measuring diet. Texas A & M Distinguished Lecturer Series. <http://stat.tamu.edu/~carroll/talks.php>
- Carroll, R. J., Ruppert, D. and Stefanski, L. A. 2006. *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd edn. Chapman and Hall/CRC.
- Chalmers, I. and Altman, D. G. 1995. *Systematic Reviews*. BMJ Publishing Group.

- Chambers, J. M. 2007. *Software for Data Analysis: Programming with R*. Springer.
- Chanter, D. O. 1981. The use and misuse of regression methods in crop modelling. In D. A. Rose and D. A. Charles-Edwards (eds), *Mathematics and Plant Physiology*. Academic Press.
- Chatfield, C. 2002. Confessions of a statistician. *The Statistician* 51: 1–20.
- Chatfield, C. 2003a. *The Analysis of Time Series: An Introduction*, 6th edn. Chapman and Hall.
- Chatfield, C. 2003b. *Problem Solving. A Statistician's Guide*, 2nd edn. Chapman and Hall/CRC.
- Clarke, D. 1968. *Analytical Archaeology*. Methuen.
- Cleveland, W. S. 1981. Lowess: a program for smoothing scatterplots by robust locally weighted regression. *The American Statistician* 35: 54.
- Cleveland, W. S. 1993. *Visualizing Data*. Hobart Press.
- Cleveland, W. S. 1994. *The Elements of Graphing Data*, revised edn. Hobart Press.
- Cochran, W. G. and Cox, G. M. 1957. *Experimental Designs*, 2nd edn. John Wiley.
- Collett, D. 2003. *Modelling Survival Data in Medical Research*, 2nd edn. Chapman and Hall.
- Cook, D. and Swayne, D. F. 2007. *Interactive and Dynamic Graphics for Data Analysis*. Springer.
- Cook, R. D. and Weisberg, S. 1999. *Applied Regression Including Computing and Graphics*. John Wiley.
- Cox, D. R. 1958. *Planning of Experiments*. John Wiley.
- Cox, D. R. and Reid, N. 2000. *Theory of the Design of Experiments*. Chapman and Hall.
- Cox, D. R. and Wermuth, N. 1996. *Multivariate Dependencies: Models, Analysis and Interpretation*. Chapman and Hall.
- Cox, T. F. and Cox, M. A. A. 2001. *Multidimensional Scaling*, 2nd edn. Chapman and Hall.
- Dalgaard, P. 2008. *Introductory Statistics with R*, 2nd edn. Springer.
- Davison, A. C. and Hinkley, D. V. 1997. *Bootstrap Methods and Their Application*. Cambridge University Press.
- Dehejia, R. H. and Wahba, S. 1999. Causal effects in non-experimental studies: re-evaluating the evaluation of training programs. *Journal of the American Statistical Association* 94: 1053–62.
- Diggle, P. 1990. *Time Series: A Biostatistical Introduction*. Clarendon Press.
- Diggle, P. J., Heagerty, P. J., Liang, K.-Y. and Zeger, S. L. 2002. *Analysis of Longitudinal Data*, 2nd edn. Clarendon Press.
- Dobson, A. J. 2001. *An Introduction to Generalized Linear Models*, 2nd edn. Chapman and Hall.
- Donner, A. and Klar, N. 2000. *Design and Analysis of Cluster Randomization Trials in Health Research*. Edward Arnold.
- Durbin, R. S., Eddy, A., Krogh, A. and Mitchison, G. 1998. *Biological Sequence Analysis*. Cambridge University Press.
- Edwards, D. 2000. *Introduction to Graphical Modelling*, 2nd edn. Springer.
- Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. 2003. Least angle regression. http://www-stat.stanford.edu/~hastie/Papers/LARS/LeastAngle_2002.pdf
- Efron, B. and Tibshirani, R. 1993. *An Introduction to the Bootstrap*. Chapman and Hall.
- Eubank, R. L. 1999. *Nonparametric Regression and Spline Smoothing*, 2nd edn. Marcel Dekker.
- Ewens, W. J. and Grant, G. R. 2005. *Statistical Methods in Bioinformatics: an Introduction*, 2nd edn. Springer.
- Fan, J. and Gijbels, I. 1996. *Local Polynomial Modelling and Its Applications*. Chapman and Hall.
- Faraway, J. J. 2004. *Linear Models with R*. Chapman and Hall/CRC.
- Faraway, J. J. 2006. *Extending the Linear Model with R. Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman and Hall/CRC.
- Finney, D. J. 1978. *Statistical Methods in Bioassay*, 3rd edn. Macmillan.
- Fisher, R. A. 1935. *The Design of Experiments*. Oliver and Boyd (7th edn, 1960).
- Fox, J. 2002. *An R and S-PLUS Companion to Applied Regression*. Sage Books.

- Gardner, M. J., Altman, D. G., Jones, D. R. and Machin, D. 1983. Is the statistical assessment of papers submitted to the *British Medical Journal* effective? *British Medical Journal* 286: 1485–8.
- Gaver, D. P., Draper, D. P., Goel, K. P., Greenhouse, J. B., Hedges, L. V., Morris, C. N. and Waternaux, C. 1992. *Combining Information: Statistical Issues and Opportunities for Research*. National Research Council, National Academy Press.
- Gelman, A. and Hill, J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gelman, A. B., Carlin, J. S., Stern, H. S. and Rubin, D. B. 2003. *Bayesian Data Analysis*, 2nd edn. Chapman and Hall/CRC.
- Gentleman, R. 2008. *R Programming for Bioinformatics*. Chapman and Hall/CRC.
- Gentleman, R. and Lang, D. 2004. Statistical analyses and reproducible research. Bioconductor Project Working Papers. Working Paper 2. <http://www.bepress.com/bioconductor/paper2>
- Gentleman, R., Carey, V., Huber, W., Irizarry, R. and Dudoit, S. 2005. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer.
- Gigerenzer, G. 1998. We need statistical thinking, not statistical rituals. *Behavioural and Brain Sciences* 21: 199–200.
- Gigerenzer, G. 2002. *Reckoning with Risk: Learning to Live with Uncertainty*. Penguin Books.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J. and Krüger, L. 1989. *The Empire of Chance*. Cambridge University Press.
- Gill, J. 2008. *Bayesian Methods: A Social and Behavioral Sciences Approach*, 2nd edn. Chapman and Hall/CRC.
- Goldstein, H. 1995. *Multilevel Statistical Models*. Arnold. <http://www.arnoldpublishers.com/support/goldstein.htm>
- Gordon, A. D. 1999. *Classification*, 2nd edn. Chapman and Hall/CRC.
- Gourieroux, C. 1997. *ARCH Models and Financial Applications*. Springer.
- Hall, P. 2001. Biometrika centenary: nonparametrics. *Biometrika* 88: 143–65.
- Harlow, L. L., Mulaik, S. A. and Steiger, J. H. (eds) 1997. *What If There Were No Significance Tests?* Lawrence Erlbaum Associates.
- Harrell, F. E. 2001. *Regression Modelling Strategies, with Applications to Linear Models, Logistic Regression and Survival Analysis*. Springer.
- Hastie, T., Tibshirani, R. and Friedman, J. 2009. *The Elements of Statistical Learning. Data Mining, Inference and Prediction*, 2nd edn. Springer.
- Hauck, W. W. J. and Donner, A. 1977. Wald's test as applied to hypotheses in logit analysis. *Journal of the American Statistical Association* 72: 851–3.
- Hoaglin, D. C. 2003. John W. Tukey and data analysis. *Statistical Science* 18: 311–18.
- Hyndman, R. J. and Khandakar, Y. 2008. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software* 27(3): 1–22. <http://www.jstatsoft.org/v27/i03>
- Hyndman, R. J., Koehler, A. B., Snyder, R. D. and Grose, S. 2002. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* 18: 439–54.
- Hyndman, R. J., Koehler, A. B., Ord, J. K. and Snyder, R. D. 2008. *Forecasting with Exponential Smoothing: The State Space Approach*. Springer.
- Ihaka, R. and Gentleman, R. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5: 299–314.
- Izenman, A. J. 2008. *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*. Springer.

- Jackson, R., Broad, J., Connor, J. and Wells, S. 2005. Alcohol and ischaemic heart disease: probably no free lunch. *The Lancet* 366: 1911–12.
- Johnson, D. H. 1995. Statistical sirens: the allure of nonparametrics. *Ecology* 76: 1998–2000.
- Krantz, D. H. 1999. The null hypothesis testing controversy in psychology. *Journal of the American Statistical Association* 44: 1372–81.
- Krzanowski, W. J. 2000. *Principles of Multivariate Analysis. A User's Perspective*, revised edn. Clarendon Press.
- Leavitt, S. D. and Dubner, S. J. 2005. *Freakonomics. A Rogue Economist Explores the Hidden Side of Everything*. William Morrow.
- Leek, J. T. and Storey, J. D. 2007. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* 3(9): e161. doi: 10.1371/journal.pgen.0030161.
- Leisch, F. 2002. *Sweave User Manual*. <http://www.ci.tuwien.ac.at/~leisch/Sweave>
- Liaw, A. and Wiener, M. 2002. Classification and regression by randomforest. *R News* 2(3): 18–22.
- Lim, T.-S. and Loh, W.-Y. 2000. A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning* 40: 203–28.
- Lumley, T. 2004a. Programmers' niche: a simple class, in S3 and S4. *R News* 4(1): 33–6.
- Lumley, T. 2004b. The survival package. *R News* 4(1): 26–8.
- Maindonald, J. H. 1984. *Statistical Computation*. John Wiley.
- Maindonald, J. H. 1992. Statistical design, analysis and presentation issues. *New Zealand Journal of Agricultural Research* 35: 121–41.
- Maindonald, J. H. 2003. The role of models in predictive validation. Invited Paper.
- Maindonald, J. 2006. Data mining methodological weaknesses and suggested fixes. In P. Christen, P. J. Kennedy, L. Jiuyong, S. J. Simoff and G. J. Williams (eds), *Fifth Australasian Data Mining Conference (AusDM2006)*, volume 61 of *CRPIT*, pp. 9–16. ACS, Sydney, Australia. <http://crpit.com/abstracts/CRPITV61Maindonald.html>
- Maindonald, J. H. 2008. Using R for data analysis and graphics. <http://www.maths.anu.edu.au/~johnm/r/usingR.pdf>
- Maindonald, J. H. and Burden, C. J. 2005. Selection bias in plots of microarray or other data that have been sampled from a high-dimensional space. In *Proceedings of 12th Computational Techniques and Applications Conference CTAC-2004*, volume 46, pp. C59–C74. <http://www.maths.anu.edu.au/~johnm/dm/ausdm06/ausdm06-jm.pdf>
- Maindonald, J. H. and Cox, N. R. 1984. Use of statistical evidence in some recent issues of DSIR agricultural journals. *New Zealand Journal of Agricultural Research* 27: 597–610.
- Maindonald, J. H., Waddell, B. C. and Petry, R. J. 2001. Apple cultivar effects on codling moth (Lepidoptera: Tortricidae) egg mortality following fumigation with methyl bromide. *Postharvest Biology and Technology* 22: 99–110.
- Manly, B. F. J. 2005. *Multivariate Statistical Methods. A Primer*, 3rd edn. Chapman & Hall/CRC.
- McCullagh, P. and Nelder, J. A. 1989. *Generalized Linear Models*, 2nd edn. Chapman and Hall.
- Meyer, D. 2001. Support vector machines. *R News* 1(3): 23–6.
- Miller, R. G. 1986. *Beyond ANOVA, Basics of Applied Statistics*. John Wiley.
- Muenchen, R. A. 2008. *R for SAS and SPSS Users*. Springer.
- Murrell, P. 2005. *R Graphics*. Chapman and Hall/CRC. <http://www.stat.auckland.ac.nz/~paul/RGraphics/rgraphics.html>
- Myers, R. H. 1990. *Classical and Modern Regression with Applications*, 2nd edn. Brooks Cole.
- Nelder, J. A. 1999. From statistics to statistical science. *Journal of the Royal Statistical Society, Series D* 48: 257–67.
- Nicholls, N. 2000. The insignificance of significance testing. *Bulletin of the American Meteorological Society* 81: 981–6.

- Ord, J. K., Koehler, A. B. and Snyder, R. D. 1997. Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association* 92: 1621–9.
- Paradis, E. 2006. *Analysis of Phylogenetics and Evolution with R*. Springer.
- Payne, R. W., Lane, P. W., Digby, P. G. N., Harding, S. A., Leech, P. K., Morgan, G. W., Todd, A. D., Thompson, R., Tunnicliffe Wilson, G., Welham, S. J. and White, R. P. 1997. *Genstat 5 Release 3 Reference Manual*. Oxford University Press.
- Pinheiro, J. C. and Bates, D. M. 2000. *Mixed Effects Models in S and S-PLUS*. Springer.
- R Development Core Team. 2009a. An introduction to R. The most recent version is available from CRAN sites. <http://cran.r-project.org>
- R Development Core Team. 2009b. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
<http://www.R-project.org>
- R Development Core Team. 2009c. *R Language Definition*. Available from CRAN sites.
- Ramsay, J. and Silverman, B. 2002. *Applied Functional Data Analysis*. Springer.
- Rao, C. and Wu, Y. 2001. On model selection (with discussion). In P. Lahiri (ed.), *Model Selection* volume 38 of *IMS Lecture Notes – Monograph Series*, pp. 1–64. Institute of Mathematical Statistics.
- Ripley, B. D. 1996. *Pattern Recognition and Neural Networks*. Cambridge University Press.
- Rosenbaum, P. R. 1999. Choice as an alternative to control in observational studies. *Statistical Science* 14: 259–78. With following discussion, pp. 279–304.
- Rosenbaum, P. R. 2002. *Observational Studies*, 2nd edn. Springer.
- Rosenbaum, P. R. 2005. Reasons for effects. *Chance* 18: 5–10.
- Rosenbaum, P. and Rubin, D. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.
- Sammon, J. W. 1969. A non-linear mapping for data structure analysis. *IEEE Transactions on Computers* C-18: 401–9.
- Sarkar, D. 2002. Lattice. *R News* 2(2): 19–23.
- Sarkar, D. 2007. *Lattice: Multivariate Data Visualization with R*. Springer. <http://lmdvr.r-forge.r-project.org/figures/figures.html>
- Schatzkin, A., Kipnis, V., Carroll, R., Midthune, D., Subar, A., Bingham, S., Schoeller, D., Troiano, R. and Freedman, L. 2003. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based observing protein and energy nutrition (open) study. *International Journal of Epidemiology* 32: 1054–62.
- Schmidt-Nielsen, K. 1984. *Scaling. Why Is Animal Size So Important?* Cambridge University Press.
- Senn, S. 2003. *Dicing with Death: Chance, Risk and Health*. Cambridge University Press.
- Sharp, S. J., Thompson, S. G. and Altman, D. G. 1996. The relation between treatment benefit and underlying risk in meta-analysis. *British Medical Journal* 313: 735–8.
- Shumway, R. and Stoffer, D. 2006. *Time Series Analysis and Its Applications: With R Examples*. Springer.
- Simpson, E. H. 1951. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B* 13: 238–41.
- Smyth, G. K. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* 3 (1), Article 3.
- Snijders, T. A. B. and Bosker, R. J. 1999. *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modelling*. Sage Books.
- Spector, P. 2008. *Data Manipulation with R*. Springer.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 64: 583–616. With following discussion, pp. 616–39.

- Sprent, P. 1966. A generalized least squares approach to linear functional relationships. *Journal of the Royal Statistical Society, Series B* 28: 278–88. With following discussion, pp. 288–97.
- Steel, R. G. D., Torrie, J. H. and Dickie, D. A. 1993. *Principles and Procedures of Statistics, A Biometrical Approach*, 3rd edn. McGraw-Hill.
- Stidd, C. K. 1953. Cube-root-normal precipitation distributions. *Transactions of the American Geophysical Union* 34: 31–5.
- Streiner, D. L. and Norman, G. R. 2003. *Health Measurement Scales. A Practical Guide to their Development and Use*, 3rd edn. Oxford University Press.
- Talbot, M. 1984. Yield variability of crop varieties in the U.K. *Journal of the Agricultural Society of Cambridge* 102: 315–21.
- Therneau, T. M. and Atkinson, E. J. 1997. An introduction to recursive partitioning using the rpart routines. Technical Report 61, Department of Health Science Research, Mayo Clinic, Rochester, MN. <http://mayoresearch.mayo.edu/mayo/research/biostat/techreports.cfm>
- Therneau, T. M. and Grambsch, P. M. 2001. *Modeling Survival Data: Extending the Cox Model*. Springer.
- Tufte, E. R. 1997. *Visual Explanations*. Graphics Press.
- Tukey, J. W. 1991. The philosophy of multiple comparisons. *Statistical Science* 6: 100–16.
- Turner, R. M., Spiegelhalter, D. J., Smith, G. C. S. and Thompson, S. G. 2009. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society, Series A* 172(1): 21–47. <http://ideas.repec.org/a/bla/jorssa/v172y2009i1p21-47.html>
- Vaida, F. and Blanchard, S. 2005. Conditional Akaike information for mixed-effects models. *Biometrika* 92: 351–70.
- Venables, W. N. 1998. Exegeses on linear models. In *Proceedings of the 1998 International S-PLUS User Conference*. <http://www.stats.ox.ac.uk/pub/MASS3/Compl.html>
- Venables, W. N. and Ripley, B. D. 2002. *Modern Applied Statistics with S*, 4th edn. Springer. See also R Complements to Modern Applied Statistics with S. <http://www.stats.ox.ac.uk/pub/MASS4/>
- Wainer, H. 1997. *Visual Revelations*. Springer.
- Weisberg, S. 1985. *Applied Linear Regression*, 2nd edn. John Wiley.
- Welch, B. L. 1949. Further note on Mrs. Aspin's tables and on certain approximations to the tabled function. *Biometrika* 36: 293–6.
- Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer. <http://had.co.nz/ggplot2/>
- Wilkinson, L. 2005. *The Grammar of Graphics*. Springer.
- Wilkinson, L. and Task Force on Statistical Inference. 1999. Statistical methods in psychology journals: guidelines and explanation. *American Psychologist* 54: 594–604.
- Williams, E. R., Matheson, A. C. and Harwood, C. E. 2002. *Experimental Design and Analysis for Use in Tree Improvement*. CSIRO Information Services. Revised.
- Williams, G. P. 1983. Improper use of regression equations in the earth sciences. *Geology* 11: 195–7.
- Wonnacott, T. H. and Wonnacott, R. 1990. *Introductory Statistics*, 5th edn. John Wiley.
- Wood, S. N. 2001. mgcv: GAMs and generalized ridge regression for R. *R News* 1(2): 20–25.
- Wood, S. N. 2006. *Generalized Additive Models. An Introduction with R*. Chapman and Hall/CRC.
- Würtz, D. 2004. *Rmetrics: An Environment for Teaching Financial Engineering and Computational Finance with R*. Rmetrics, ITP, ETH Zürich, Switzerland. <http://www.rmetrics.org>
- Xie, Y. and Cheng, X. 2008. animation: A package for statistical animations. *R News* 8(2): 23–7.

- Young, G. and Smith, R. L. 2005. *Essentials of Statistical Inference*. Cambridge University Press.
- Zeger, S. L., Thomas, D., Dominici, F., Samet, J. M., Schwartz, J., Dockery, D. and Cohen, A. 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environmental Health Perspectives* 108: 419–26. See also vol. 109, p. A517.
- Zhang, H. and Singer, B. 1999. *Recursive Partitioning in the Health Sciences*. Springer.
- Zhu, X., Ambroise, C. and McLachlan, G. J. 2006. Selection bias in working with the top genes in supervised classification of tissue samples. *Statistical Methodology* 3: 29–41.

References for Data Sets

- Andrews, D. F. and Herzberg, A. M. 1985. *Data. A Collection of Problems from Many Fields for the Student and Research Worker*. Springer.
- Ash, J. and Helman, C. 1990. Floristics and vegetation biomass of a forest catchment, Kioloa, south coastal N.S.W. *Cunninghamia* 2: 167–82.
- Blake, C. and Merz, C. 1998. UCI repository of machine learning databases. <http://www.ics.uci.edu/~mllearn/MLRepository.html>
- Boot, H. M. and Maindonald, J. 2008. New estimates of age- and sex-specific earnings and the male–female earnings gap in the British cotton industry, 1833–1906. *Economic History Review* 61: 380–408.
- Burns, N. R., Nettlebeck, T., White, M. and Willson, J. 1999. Effects of car window tinting on visual performance: a comparison of elderly and young drivers. *Ergonomics* 42: 428–43.
- Bussolari, S. 1987. Human factors of long-distance human-powered aircraft flights. *Human Power* 5: 8–12.
- Charig, C. R. 1986. Comparison of treatment of renal calculi by operative surgery, percutaneous nephrolithotomy, and extracorporeal shock wave lithotripsy. *British Medical Journal* 292: 879–82.
- Christie, M. 2000. *The Ozone Layer: A Philosophy of Science Perspective*. Cambridge University Press.
- Chu, I., Secours, V., Villeneuve, D. C., Valli, V. E., Nakamura, A., Colin, D., Clegg, D. J. and Arnold, E. P. 1988. Reproduction study of toxaphene in the rat. *Journal of Environmental Science and Health Part B. Pesticides and Food Contamination* 23: 101–26.
- Clutton-Brock, T. H., O’Riain, M. J., Brotherton, P. N. M., Gaynor, D., Kansky, R., Griffin, A. S. and Manser, M. 1999. Selfish sentinels in cooperative mammals. *Science*, pp. 1640–44.
- Cohen, P. 1996. Pain discriminates between the sexes. *New Scientist* 2054: 16.
- Daniels, M., Devlin, B. and Roeder, K. 1997. Of genes and IQ. In B. Devlin, S. Fienberg and K. Roeder (eds), *Intelligence, Genes and Success*, Chapter 3. Springer.
- Darwin, C. 1877. *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*. Appleton and Company.
- Dehejia, R. H. and Wahba, S. 1999. Causal effects in non-experimental studies: re-evaluating the evaluation of training programs. *Journal of the American Statistical Association* 94: 1053–62.
- Ezzet, F. and Whitehead, J. 1991. A random effects model for ordinal responses from a crossover trial. *Statistics in Medicine* 10: 901–7.
- Farmer, C. 2005. Another look at Meyer and Finney’s ‘who wants airbags?’ *Chance* 19: 15–22.
- Gahr, M. and Pilleri, G. 1969. Anatomy and biometry of *Stenella* and *Delphinus*. In G. Pilleri (ed.), *Investigations on Cetacea*. Hirnanatomisches Institute der Universität Bern.
- Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D. and Lander, E. S. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531–7.

- Gordon, N. C., Gear, R. W., Heller, P. H. and Levine, J. D. 1995. Enhancement of morphine analgesia by the GABA[subscript B] agonist baclofen. *Neuroscience* 69: 345–9.
- Gordon, W. 1894. *Our Country's Birds and How to Know Them*. Day and Son.
- Grasso, L. C., Maindonald, J., Rudd, S., Hayward, D. C., Saint, R., Miller, D. J. and Ball, E. E. 2008. Microarray analysis identifies candidate genes for key roles in coral development. *BMC Genomics* 9: 540.
- Guy, W. A. 1882. Two hundred and fifty years of small pox in London, together with a supplement relating to England and Wales. *Journal of the Royal Statistical Society* 45: 399–443.
- Hales, S., de Wet, N., Maindonald, J. and Woodward, A. 2002. Potential effect of population and climate change global, distribution of dengue fever: an empirical model. *The Lancet* 360: 830–34.
- Hall, P. 2003. A possum's tale – how statistics revealed a new mammal species. *Chance* 16: 8–13.
- Harker, F. R. and Maindonald, J. H. 1994. Ripening of nectarine fruit. *Plant Physiology* 106: 165–71.
- Hobson, J. A. 1988. *The Dreaming Brain*. Basic Books.
- Hunter, D. 2000. *The conservation and demography of the southern corroboree frog (Pseudophryne corroboree)*. MSc, University of Canberra.
- Jouzel, J. *et al.* 2007. EPICA Dome C ice core 800kyr deuterium data and temperature estimates. IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series # 2007-091. NOAA/NCDC Paleoclimatology Program, Boulder, CO, USA.
- King, D. A. 1998. Relationship between crown architecture and branch orientation in rain forest trees. *Annals of Botany* 82: 1–7.
- King, D. A. and Maindonald, J. H. 1999. Tree architecture in relation to leaf dimensions and tree stature in temperate and tropical rain forests. *Journal of Ecology* 87: 1012–24.
- Lalonde, R. 1986. Evaluating the economic evaluations of training programs. *American Economic Review* 76: 604–20.
- Latter, O. H. 1902. The egg of *cuculus canorus*. An inquiry into the dimensions of the cuckoo's egg and the relation of the variations to the size of the eggs of the foster-parent, with notes on coloration, &c. *Biometrika* 1: 164–76.
- Linacre, E. 1992. *Climate Data and Resources. A Reference and Guide*. Routledge.
- Linacre, E. and Geerts, B. 1997. *Climates and Weather Explained*. Routledge.
- Linde, K., Streng, A., Jurgens, S., Hoppe, A., Brinkhaus, B., Witt, C., Wagenpfeil, S., Pfaffenrath, V., Hammes, M., Weidenhammer, W., Willich, S. and Melchart, D. 2005. Acupuncture for patients with migraine. A randomized controlled trial. *Journal of the American Medical Association* 293: 2118–25.
- Lindenmayer, D. B., Viggers, K. L., Cunningham, R. B. and Donnelly, C. F. 1995. Morphological variation among columns of the mountain brushtail possum, *Trichosurus caninus* Ogilby (Phalangeridae: Marsupiala). *Australian Journal of Zoology* 43: 449–58.
- Lüthi, D., Floch, M. L., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K. and Stocker, T. 2008. High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* 453: 379–82.
- Marland, G., Boden, T. A. and Andres, R. J. 2003. Global, regional, and national CO₂ emissions. In *Trends: A Compendium of Data on Global Change*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, USA. <http://cdiac.esd.ornl.gov/>
- Matthews, D. E. and Farewell, V. T. 1996. *Using and Understanding Medical Statistics*. Karger.
- McLellan, E. A., Medline, A. and Bird, R. P. 1991. Dose response and proliferative characteristics of aberrant crypt foci: putative preneoplastic lesions in rat colon. *Carcinogenesis* 12: 2093–8.
- McLeod, C. C. 1982. Effect of rates of seeding on barley grown for grain. *New Zealand Journal of Agriculture* 10: 133–6.

- Meyer, M. 2006. Commentary on “another look at Meyer and Finney’s ‘who wants airbags?’” *Chance* 19: 23–2.
- Meyer, M. and Finney, T. 2005. Who wants airbags? *Chance* 18: 3–16.
- Mitchell, B. R. 1988. *British Historical Statistics*. Cambridge University Press.
- Nadel, E. and Bussolari, S. 1988. The Daedalus project: physiological problems and solutions. *American Scientist* 76: 351–60.
- Newton, A. 1893–1896. Cuckoos. In A. Newton and H. Gadow (eds), *Dictionary of Birds*. A. & C. Black.
- Nicholls, N., Lavery, B., Frederiksen, C. and Drosowsky, W. 1996. Recent apparent changes in relationships between the El Niño, southern oscillation and Australian rainfall and temperature. *Geophysical Research Letters* 23: 3357–60.
- Nightingale, F. 1871. *Notes on Lying-in Institutions*. Longmans, Green and Co.
- Perrine, F. M., Prayitno, J., Weinman, J. J., Dazzo, F. B. and Rolfe, B. 2001. Rhizobium plasmids are involved in the inhibition or stimulation of rice growth and development. *Australian Journal of Plant Physiology* 28: 923–7.
- Roberts, H. V. 1974. *Conversational Statistics*. Hewlett-Packard University Business Series. The Scientific Press.
- Shanklin, J. 2001. Ozone at Halley, Rothera and Vernadsky/Faraday. www.antarctica.ac.uk/met/jds/ozone/
- Sharp, W. D. and Clague, D. A. 2006. 50-yr initiation of Hawaiian–Emperor bend records major change in Pacific Plate motion. *Science* 313: 1281–4.
- Snelgar, W. P., Manson, P. J. and Martin, P. J. 1992. Influence of time of shading on flowering and yield of kiwifruit vines. *Journal of Horticultural Science* 67: 481–7.
- Stewardson, C. L., Hemsley, S., Meyer, M. A., Canfield, P. J. and Maindonald, J. H. 1999. Gross and microscopic visceral anatomy of the male Cape fur seal, *Arctocephalus pusillus pusillus* (Pinnipedia: Otariidae), with reference to organ size and growth. *Journal of Anatomy (Cambridge)* 195: 235–55. (WWF project ZA-348.)
- Stewart, K. M., Van Toor, R. F. and Crosbie, S. F. 1988. Control of grass grub (Coleoptera: Scarabaeidae) with rollers of different design. *New Zealand Journal of Experimental Agriculture* 16: 141–50.
- Stiell, I. G., Wells, G. A., Vandemheen, K., Clement, C., Lesiuk, H., Laupacis, A., McKnight, R. D., Verbeek, R., Brison, R., Cass, D., Eisenhauer, M. A., Greenberg, G. H. and Worthington, J. F. 2001. The Canadian CT head rule for patients with minor head injury. *The Lancet* 357: 1391–6.
- Stocks, P. 1942. Measles and whooping cough during the dispersal of 1939–1940. *Journal of the Royal Statistical Society* 105: 259–91.
- Telford, R. D. and Cunningham, R. B. 1991. Sex, sport and body-size dependency of hematology in highly trained athletes. *Medicine and Science in Sports and Exercise* 23: 788–794.
- Thall, P. F. and Vail, S. C. 1990. Some covariance models for longitudinal count data. *Biometrics* 46: 657–71.
- Tippett, L. H. C. 1931. *The Methods of Statistics*. Williams and Norgate.
- Wainright, P., Pelkman, C. and Wahlsten, D. 1989. The quantitative relationship between nutritional effects on preweaning growth and behavioral development in mice. *Developmental Psychobiology* 22: 183–93.

References for Packages

- base*: *base*, *datasets*, *grDevices*, *graphics*, *grid*, *methods*, *splines*, *stats*, *stats4*, *tcltk*, *tools*, *utils*,
 R Development Core Team, 2008. R Foundation for Statistical Computing. <http://www.r-project.org>

- boot: Bootstrap R (S-Plus) Functions* (Canty), Canty, A. and Ripley, B. 2008. (Version 1.2-34.) See further Canty (2002).
- car: Companion to Applied Regression*, Fox, J. 2008. (V. 1.2-8.) <http://socserv.socsci.mcmaster.ca/jfox/>
- cluster: Cluster Analysis Basics and Extensions*, Maechler, M., Rousseeuw, P., Struyf, A. and Hubert, M. 2008. (V. 1.11.11.) Based on S original by P. Rousseeuw, A. Struyf and M. Hubert; initial R port by K. Hornik.
- colorRamps: Builds color tables*, Keitt, T. 2007. (V. 2.2.)
- compositions: Compositional Data Analysis*, van den Boogaart, K. G., 2009. (V. 0.91-6.)
- DAAG: Data Analysis And Graphics*, Maindonald, J. and Braun, W. J. 2009. (V. 0.98.) <http://www.stats.uwo.ca/DAAG>
- DAAGxtras: Data sets and Functions, supplementary to DAAG*, Maindonald, J. 2009. (R package version 0.7-7.) <http://www.maths.anu.edu.au/~johnm>
- dichromat: Color schemes for dichromats*, Lumley, T. 2007. (V. 1.2-3.)
- fgui: Automatic creation of Graphical User Interfaces for command-line R packages*, Hoffmann, T. J. 2009.
- Forecasting (functions and datasets for forecasting): bundle of forecast, fma, Mcomp, expsmooth*, Hyndman, R. J. 2009. (R package version 1.24.) <http://www.robhyndman.info/Rlibrary/forecast/>
- foreign: Read Data from Minitab, S, SAS, SPSS, Stata, Systat, dBase, . . .*, R-core, DebRoy, S., Bivand, R. and others. 2008. (V. 0.8-33.) See COPYRIGHTS file in the sources for acknowledgments.
- ggplot2: An implementation of the Grammar of Graphics*, Wickham, H. 2009. (V. 0.8.3.) <http://had.co.nz/ggplot2/>
- golubEsets: exprSets for golub leukemia data*, Golub, T. 2002. (V. 1.0.1.) Maintained by V. Carey.
- hddplot: Use known groups in high-dimensional data to derive scores for plots*, Maindonald, J. H. 2006. (V. 0.5-0.) See further Maindonald and Burden (2005).
- lattice: Lattice Graphics*, Sarkar, D. 2009. (V. 0.17-20.) See further Sarkar (2002).
- latticeExtra: Extra Graphical Utilities Based on Lattice*, Sarkar, D. and Andrews, F. 2008. (V. 0.5-4.) See further Sarkar (2002).
- latticist: A Lattice-based tool for exploratory visualisation*, Andrews, F. 2008. (R package version 0.9-40.) <http://latticist.googlecode.com/>
- leaps: regression subset selection*, Lumley, T. and Miller, A. 2009. (V. 2.8.) Compiled by T. Lumley; Fortran code by A. Miller. See further Miller (1986).
- lme4: Linear mixed-effects models using S4 classes*, Bates, D., Maechler, M. and Dai, B. 2008. (V. 0.999375-28.) See further Bates (2005).
- MCMCpack: Markov chain Monte Carlo (MCMC) Package*, Martin, A. D., Quinn, K. M. and Park, J. H. 2009. (V. 0.9-6.)
- MEMSS: Data sets from Mixed-effects Models in S*, Bates, D. 2008. (V. 0.3-4.)
- mgcv: GAMs and Generalized Ridge Regression for R*, Wood, S. N. 2008. (V. 1.4-1.) See further Wood (2001).
- monoProc: strictly monotone smoothing procedure*, Scheder, R. 2005. (V. 1.0-4.)
- multtest: Resampling-based multiple hypothesis testing*, Pollard, K. S., Ge, Y. and Dudoit, S. 2005. (V. 1.8.0.)
- nlme: Linear and nonlinear mixed effects models*, Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R core team. 2008. (V. 3.1-89.) See further Pinheiro and Bates (2000).
- oz: Plot the Australian coastline and states*, Venables, W. N. and Hornik, K. 2007. (V. 1.0-16.) S original by W. Venables; R port by K. Hornik.
- playwith: A GUI for interactive plots using GTK+*, Andrews, F. 2008. (R package version 0.9-43; requires Gtk+.) <http://playwith.googlecode.com/>

- randomForest: Classification and Regression*, Liaw, A. and Wiener, M. 2008. (V. 4.5-30.) See further Liaw and Wiener (2002).
- rattle: A graphical user interface for data mining in R using GTK*, Williams, G. 2009. (R package version 2.4.54.) <http://rattle.togaware.com/>
- RColorBrewer: ColorBrewer palettes*, Neuwirth, E. 2007. (V. 1.0-2.)
- reshape: Flexibly reshape data*, Wickham, H. 2008. (V. 0.8.2.) <http://had.co.nz/reshape>
- rpart: Recursive Partitioning*, Therneau, T. M. and Atkinson, B. 2008. (V. 3.1-42.) R port by B. Ripley.
- survival: Survival analysis, including penalised likelihood.*, Therneau, T. and Lumley, T. 2008. (V. 2.34-1.) S original by T. Therneau; R port by T. Lumley. See further Therneau and Grambsch (2001) and Lumley (2004b).
- tseries: Time series analysis and computational finance*, Trapletti, A. and Hornik, K. 2008. (V. 0.10-16.) Compiled by A. Trapletti.
- VR packages: bundle of MASS, class, nnet, spatial*, Venables, W. N. and Ripley, B. D. 2008. (V. 7.2-44.) R port by B. Ripley, following earlier work by K. Hornik and A. Gebhardt. See further Venables and Ripley (2002). <http://www.stats.ox.ac.uk/pub/MASS4/>
- xtable: Export tables to LaTeX or HTML*, Dahl, D. and others. 2008. (V. 1.5-4.)
- Other packages and package bundles mentioned in the text are:* *ape* (Paradis, E., Claude, J. and Strimmer, K.), *Deducer* (GUI; Fellows, I.), *Devore6* (Devore, J. L. and Bates, D.), *dr* (Weisberg, S.), *dse* (package bundle; Gilbert, P.), *fseries* (Wuertz, D.), *gam* (Hastie, T.), *JGR* (Helbig, H. and Urbanek, S.), *lars* (Hastie, T. and Efron, B.), *KernSmooth* (Wand, M. and Ripley, B.), *locfit* (Loader, C.), *MPV* (Braun, W. J.; data sets from Montgomery, Peck and Vining), *muhaaz* (Hess, K. and Gentleman, R.), *plyr* (Wickham, H.), *pmg* (Verzani, J.), *RODBC* (Lapsley, M. and Ripley, B. D.) and *strucchange* (Zeileis *et al.*). For more complete acknowledgments, see the help pages for the individual packages. See also the information given by calling `citation()` with the package name as character string argument.

References for Web Pages

Comprehensive R Archive Network.

<http://cran.r-project.org/>

CRAN Task Views.

<http://cran.r-project.org/web/views/>

ESS – Emacs Speaks Statistics.

<http://ess.r-project.org/>

fgui: Rapidly create a gui interface, Hoffmann, T. J. and Laird, N. M. 2009 (V.1.0-0)

<http://www.people.fas.harvard.edu/~tjhoffm/fgui.html>

Gtk+. The development version is required for playwith; suggested for rattle.

<http://downloads.sourceforge.net/gladewin32/>

Machine Learning Repository.

<http://www.ics.uci.edu/~mlearn/MLRepository.html>

R News.

<http://cran.r-project.org/doc/Rnews/>

S+, Tibco Spotfire S+ (formerly S-PLUS).

<http://spotfire.tibco.com/>

Tinn-R.

<http://sourceforge.net/projects/tinn-r>

Acknowledgments for use of data

We thank the following for permission to reproduce graphs or tables that have appeared in published material: *Journal of Ecology*, for permission to reproduce Figure 12.5B, which is a redrawn version of the fourth panel in Figure 3 in King and Maindonald (1999); *Plant Physiology*, in relation to the right panel of Figure 2.6, which is a redrawn version of Figure 3 in Harker and Maindonald (1994), copyrighted by the American Society of Plant Biologists; SIR Publishing (The Royal Society of New Zealand), for permission to reproduce data in Subsection 3.1.1 from Stewart *et al.* (1988), Table 10.2 from Maindonald (1992), and Figure 10.4 which is similar to Figure 1 in Maindonald (1992); CSIRO Publishing, for permission to reproduce (in Section 4.4) data that appear in Table 4 in Perrine *et al.* (2001); *Australian Journal of Zoology*, for permission to reproduce a graph that is similar to a part of Figure 2 in Lindenmayer *et al.* (1995); *American Medical Association* (©2005, All rights reserved), for permission to use the data in Exercise 4.11 in Section 9, from Linde *et al.* (2005).

We acknowledge the help of the following individuals and organizations in making data available. Darren Kriticos (CSIRO Division of Entomology): data frame houseprices. J. D. Shanklin (British Antarctic Survey): ozone (Shanklin, 2001). W. S. Snelgar (HortResearch NZ): kiwishade (Snelgar *et al.*, 1992). Francine Adams and supervisors Rosemary Martin and Murali Nayadu (all from ANU): science. Jasmyn Lynch (ANU), for the rare and endangered plant species data in Subsection 4.3.1. D. B. Lindenmayer and coworkers (ANU): possum (Lindenmayer *et al.*, 1995). D. A. King (formerly ANU): leafshape and leafshape17 (King and Maindonald, 1999). R. F. Harker (HortResearch NZ): fruitohms (Harker and Maindonald, 1994). E. Linacre (ANU): dewpoint (Linacre, 1992, Linacre and Geerts, 1997). N. R. Burns (University of Adelaide): tinting (Burns *et al.*, 1999). M. Boot (ANU): wages1833 (Boot and Maindonald, 2008). J. Erickson (University of Chicago) and A. H. Welsh (University of Southampton): anesthetic. D. Hunter (University of Canberra): frogs. Sharyn Wragg (formerly ANU): moths. Claudia Haarman (ANU), for the flow inhibition data of Exercise 1 in Section 8.9. Melissa Manning (ANU): socsupport. J. Ash (ANU): rainforest (Ash and Helman, 1990). Katharina Siebke and Susan von Cammerer (ANU): leaftemp. Ranjana Bird (University of Manitoba), for the aberrant crypt foci data in Section 3.7. Australian Bureau of Meteorology: bomregions. Note the abbreviation ANU for Australian National University.

Acknowledgments for data from web pages

(Help pages are for the DAAG package.)

bomregions: See help(bomregions) for details.

edcT and edcCO2: See help(edcT) and help(edcCO2) for details. See Jouzel *et al.* (2007), Lüthi *et al.* (2008). The help pages for edcCO2 and edcT give additional references and links.

hills2000: <http://www.hillrunning.co.uk>

hotspots, hotspots2006: See help(hotspots); Sharp and Clague (2006).

monica: <http://www.ktl.fi/monica>

nihills: <http://www.nimra.org.uk/calendar.asp>

ozone: See help(ozone); Shanklin (2001).

nswdemo, cps1, psid1, nswpsid1: See help(nswdemo); Dehejia and Wahba (1999), Lalonde (1986).