

## References

- Aitkin,M., Anderson,D. and Hinde,J. (1981). Statistical modelling of data on teaching styles (with discussion). J. Royal Statist. Soc., A, 144, 148-61
- Aitkin,M. and Longford, N. (1986). Statistical modelling in school effectiveness studies (with discussion). J. Royal Statist. Soc., A, 149, 1-43
- Aitkin,M., Anderson,D., Francis,B. and Hinde,J. (1989). Statistical modelling in GLIM. Oxford, Clarendon Press.
- Beale,E.M.L. and Little,R.J.A. (1975). Missing values in multivariate analysis. J. Royal Statist. Soc., B, 37, 129-45.
- Bennett,N.(1976). Teaching Styles and Pupil Progress. London, Open Books.
- Bock,R.D. (1992). Structural and nonstructural analysis of multiphasic growth. Chicago, University of Chicago (unpublished).
- Breslow,N.E. and Clayton,D.G. (1993). Approximate inference in generalised linear mixed models. J. American Statistical Association, 88, 9-25.
- Bryk,A.S., and Raudenbush,S.W. (1992). Hierarchical Linear Models. Newbury Park, Sage.
- Bryk,A.S., Raudenbush, S.W., Seltzer, M. and Congdon, R. (1988). An introduction to HLM: Computer program and user's guide (2nd. Edn.) Chicago, University of Chicago Dept. of Education.
- Burdick,R.K. and Graybill,F.A. (1988). The present status of confidence interval estimation on variance components in balanced and unbalanced random models. Communications in Statistics: theory and methods, 17, 1165-1195.
- Burstein,L., Fischer,K.H., and Miller,M.D. (1980). The multilevel effects of background on science achievement: a cross national comparison. Sociology of Education, 53, 215-225
- Clayton,D.G. (1988). The analysis of event history data: a review of progress and outstanding problems. Statistics in medicine, 7, 819-841.
- Clayton,D.G. (1991). A Monte Carlo method for Bayesian inference in frailty models. Biometrics, 47, 467-485.
- Clayton,D.G. (1992). Bayesian analysis of frailty models. Cambridge, MRC Biostatistics Unit. (unpublished).
- Cochran,W.G. (1983). Planning and analysis of observational studies. New York, Wiley.
- Cook,R.D.and Weisberg,S. (1982). Residuals and Influence in regression. London, Chapman and Hall.
- Cox,D.R. (1972).Regression models and life tables (with discussion). J. Royal Statist. Soc., B, 34, 187-220.
- Cox,D.R. and Oakes,D. (1984) Analysis of Survival Data. London, Chapman and Hall.
- Creswell,M. (1991). A multilevel Bivariate Model. In: Prosser, R., Rasbash, J. and Goldstein,H. Data Analysis with ML3. London, Institute of Education.
- Cronbach,L.J. and Webb,N. (1975). Between class and within class effects in a repeated aptitude x treatment interaction: reanalysis of a study by G.L. Anderson, J. Educational Psychology, 67, 717-24.
- Demirjian,A., La Palme,L. and Thibault,H.W. (1982). La croissance staturo-ponderale des enfants Canadien-Francais de la naissance a 36 mois. Union Med Can., 112, 153-63.
- Derbyshire,M.E. (1987). Statistical rationale for grant-related expenditure assessment (GREA) concerning personal social services. J. Royal Statist. Soc., A, 150, 309-33.
- Ecob,R. and Goldstein,H. (1983). Instrumental variable methods for the estimation of test score reliability. J. Educational Statistics, 8, 223-41.
- Efron,B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. J. Amer. Statist. Asscn., 83, 414-425.

- Efron,B. and Gong,G. (1983). A leisurely look at the Bootstrap, the Jackknife and Cross-validation. *The American Statistician*, 37, 36-48.
- Egger,P.J. (1992). Event history analysis: discrete-time models including unobserved heterogeneity, with applications to birth history data. University of Southampton, PhD thesis.
- Fuller, W.A. (1987). *Measurement Error Models*. New York, Wiley.
- Garrett,M., Fitzmaurice,M. and Laird, N. (1993). A likelihood based method for analysing longitudinal binary responses. *Biometrika*, 80, 141-51.
- Gilks,W.R., Clayton,D.G., Spiegelhalter,D.J., Best,N.G., McNeil,A.J., Sharples,L.D., and Kirby,A.J. (1993). Modelling complexity: applications of Gibbs Sampling in medicine. (With discussion). *Journal of the Royal Statistical Society, B*, 55, 39-102.
- Goldstein,H. (1976). Smoking in pregnancy: some notes on the statistical controversy. *British Journal of Preventive and Social Medicine*, 31, 13-17.
- Goldstein,H. (1979). *The Design and Analysis of Longitudinal Studies*, London, Academic Press.
- Goldstein,H. (1986). Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika*, 73, 43-56.
- Goldstein,H. (1987a). Multilevel covariance component models. *Biometrika*, 74, 430-31.
- Goldstein,H. (1987b). *Multilevel Models in Educational and Social Research*. London, Griffin.
- Goldstein,H. (1987c). The choice of constraints in correspondence analysis. *Psychometrika*, 52, 207-15.
- Goldstein,H. (1989a). Restricted unbiased iterative generalised least squares estimation. *Biometrika*, 76, 622-23.
- Goldstein,H. (1989b). Efficient prediction models for adult height. In J.M.Tanner (Ed.) *Auxology 88; Perspectives in the science of growth and development*. London, Smith Gordon.
- Goldstein,H. (1991). Nonlinear multilevel models with an application to discrete response data. *Biometrika*, 78, 45-51.
- Goldstein,H. (1992). Statistical information and the measurement of education outcomes (editorial). *J.Roy.Statist.Soc.,A*. 155, 313-15.
- Goldstein, H. and Blatchford, P. (1998). Class size and educational achievement: a review of methodology with particular reference to study design. *British Educational Research Journal* **24**: 255-268.
- Goldstein,H. and McDonald,R.P. (1987). A general model for the analysis of multilevel data. *Psychometrika*, 53, 455-67.
- Goldstein,H. and Rasbash,J. (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares. *Computational Statistics and Data Analysis*, 13, 63-71.
- Goldstein, H. and Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society, A*. **159**: 505-13.
- Goldstein,H., Rasbash,J., Yang,M., Woodhouse,G., Pan,H., Nuttall,D., and Thomas,S. (1993). A multilevel analysis of school examination results. *Oxford review of education*, 19, 425-33.
- Goldstein, H., and Healy, M.J.R. (1994). The graphical presentation of a collection of means. *J. Royal Statistical Society, A*, 158, 175-7.
- Goldstein,H., Healy,M.J.R., and Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, 13, 1643-1655.
- Greenacre,M.J. (1984). *Theory and applications of correspondence analysis*. New York, Academic Press.
- Grizzle, J.C. and Allen, D.M. (1969). An analysis of growth and dose response curves. *Biometrics*, 25, 357-61.
- Gumpertz,M.L. and Pantula,S.G. (1992). Nonlinear regression with variance components. *J. American Statistical Association*, 87, 201-9.

- Harrison,G.A. and Brush,G. (1990). On correlations between adjacent velocities and accelerations in longitudinal growth data. *Annals of Human Biology*, 17, 55-57.
- Heath,A., Jowell,R., Curtice,J., Evans,G., Field,J. and Witherspoon,S. (1991). *Understanding Political Change*. Oxford, Pergamon.
- Hedges,L.V. and Olkin,I.O. (1985). *Statistical Methods for Meta Analysis*. Orlando Florida, Academic Press.
- Holland,P.W. (1986). Statistics and causal inference. *J. American Statistical Association*, 81, 945-971.
- Jenss, R.M. and Bayley, N. (1937). A mathematical method for studying the growth of a child. *Human Biology*, 9, 556-63.
- Joreskog,K.G. and Sorbom,D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA. Abt books.
- Kreft, I.G., de Leeuw,J. and van der Leeden,R.. (1994). Comparing five different statistical packages for hierarchical linear regression: BMDP-5V, GENMOD, HLM, ML3, and VARCL. *American Statistician*, 48, (to appear).
- Laird,N.M. and Louis,T.A. (1987). Empirical Bayes confidence intervals based on bootstrap samples. *J. Amer. Statist. Asscn.*, 82, 739-57.
- Laird,N.M. and Louis,T.A. (1989). Empirical Bayes confidence intervals for a series of related experiments. *Biometrics*, 45, 481-95.
- Langford, I. and Lewis, T. (1998). Outliers in multilevel data. *Journal of the Royal Statistical Society, A*. **161**: 121-160.
- Larsen,U. and Vaupel,J.W. (1993). Hutterite fecundability by age and parity: strategies for frailty modelling of event histories. *Demography*, 30, 81-101.
- Lawley,D.N., and Maxwell,A.E. (1971). *Factor Analysis as a Statistical Method*. 2nd edition. London, Butterworth.
- Liang,K. and Zeger,S.L. (1986). Longitudinal data analysis using generalised linear models. *Biometrika*, 73, 45-51.
- Lindley,D.V. and Smith,A.F.M. (1972). Bayes estimates for the linear model. *J. Royal Statist, Soc., B*, 34, 1-41
- Lindstrom,M.J. and Bates,D.M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46,673-687.
- Little,R.J.A. (1992). Regression with missing X's: a review. *J. American Statistical Association*, 87, 1227-37.
- Longford,N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, 74, 817-27.
- Longford,N.T. (1988). VARCL - software for variance component analysis of data with hierarchically nested random effects (maximum likelihood). Princeton, N.J., Educational Testing Service.
- Longford, N.T. (1993). *Random Coefficient Models*. Oxford, Clarendon Press.
- Longford,N.T. and Muthen,B.O. (1992). Factor analysis for clustered populations. *Psychometrika*, 57, 581-97.
- McCullagh,P. and Nelder,J. (1989). *Generalised Linear Models* (2nd edition), London, Chapman and Hall.
- McDonald,R.P. (1985). *Factor Analysis and Related Methods*. Hillsdale, New York; Lawrence Erlbaum.
- McDonald,R.P. (1993). A general model for two level data with responses missing at random. *Psychometrika*, 58, 575-85.
- McDonald,R.P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological methods and research*, 22, 399-413.
- McDonald,R.P. and Goldstein,H. (1988). Balanced versus unbalanced designs for linear structural relations in two level data. *British Journal of mathematical and statistical psychology*, 42, 215-32.
- McGrath,K. and Waterton,J. (1986). *British Social Attitudes, 1983-1986 panel survey*. London, Social and Community Planning Research.

- Mason, W.M., Anderson, A.F., and Hayat, N. (1988). Manual for GENMOD. Ann Arbor, University of Michigan Population Studies Centre.
- Miller, R.G. (1974). The Jackknife - a review. *Biometrika*, 61,1-15.
- Mortimore, P., Sammons, P. Stoll, L., Lewis, D. and Ecob, R. (1988). *School Matters*. Wells, Open Books.
- Moulton, L.H. and Zeger, S.L. (1989). Analysing repeated measures on generalised linear models via the bootstrap. *Biometrics*, 45, 381-94.
- Muthen, B.O. (1989). Latent variable modelling in heterogeneous populations. *Psychometrika*, 54, 557-85.
- Nuttall, D.L., Goldstein, H., Prosser, R. and Rasbash, J. (1989). Differential school effectiveness. *International Journal of Educational Research*, 13, 769-76.
- Paterson, L. (1991). Socio-economic status and educational attainment: a multidimensional and multilevel study. *Evaluation and Research in Education*, 5, 97-121.
- Peto, R. (1972). Contribution to discussion of paper by D.R.Cox. *J. Royal Statist. Society*, B,34,205-7.
- Pfeffermann, D., Skinner, C. J., Holmes, D., Goldstein, H., et al. (1997). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society*, B **60**: 23-40.
- Plewis, I. (1985). *Analysing change*. Chichester, Wiley.
- Plewis, I. (1993). Reading Progress. In 'A Guide to ML3 for new users' ed. G. Woodhouse, London, Multilevel Models Project.
- Plewis, I. (1994). Statistical methods for understanding cognitive growth: a review, a synthesis and an application. (In press).
- Prosser, R., Rasbash, J. and Goldstein, H. (1991). *ML3 Software for Three-level analysis: user's guide for version 2*. London, Institute of Education.
- Rasbash, J., Yang, M., Woodhouse, G., and Goldstein, H. (1995). *MLn command reference*. London, Institute of Education.
- Rasbash, J. and Goldstein, H. (1994). Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *J. Educational and Behavioural Statistics*, 19, 337-50.
- Raudenbush, S.W. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *J. Educational Statistics*, 18, 321-49.
- Raudenbush, S.W. (1994). Equivalence of Fisher Scoring to Iterative Generalised Least Squares in the Normal case with application to hierarchical linear models. Unpublished.
- Robinson, W.S. (1950). Ecological correlations and the behaviour of individuals. *Amer. Sociol. Rev.*, 15, 351-57.
- Rosier, M.J. (1987). The second international science study. *Comparative Education Review* 31:106-28.
- Royall, R.M. (1986). Model robust confidence intervals using maximum likelihood estimators. *International Statistical Review*, 54, 221-26.
- Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York, Wiley.
- Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components*. New York, Wiley.
- Seltzer, M.H. (1993). Sensitivity analysis for fixed effects in the hierarchical model: a Gibbs sampling approach. *Journal of Educational Statistics*, 18,207-36.
- Skinner, C.J., Holt, D. and Smith, T.M.F (1989). *Analysis of complex surveys*, Chichester: Wiley.
- Snijders, T.A.B. and Bosker, R.J. (1993). Standard errors and sample sizes for two-level research. *J. Educational Statistics*, 18, 237-59.
- Vevea, J. (1994). A model for estimating effect size in the presence of publication bias. Paper presented to American Educational Research Association, annual meeting, New Orleans, April 1994.

- Waclawiw,M.A. and Liang,K. (1993). Prediction of random effects in the generalised linear model. J. American Statistical Association, 88, 171-78.
- Waclawiw,M.A. and Liang,K. (1994). Empirical Bayes estimation and inference for the random effects model with binary response. Statistics in Medicine, 13, 541-51.
- Wei,L.J., Lin,D.Y. and Weissfeld,L. (1989). Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. Journal of American Statistical Association, 84, 1065-73.
- Wolfinger,R. (1993). Laplace's approximation for nonlinear mixed models. Biometrika, 80, 791-795.
- Woodhouse, G. (1998). *Measurement errors in multilevel models*. PhD thesis, University of London.
- Woodhouse,G., Yang,M., Goldstein,H., Rasbash,J. and Pan,H. (1995). Adjusting for measurement error in multilevel analysis. J. of Educational and Behavioural Statistics, (to appear).
- Zeger,S.L., Liang,K-Y., and Albert,P.S. (1988). Models for longitudinal data: a generalised estimating equation approach. Biometrics, 44, 1049-60.
- Zeger,S.L. and Karim,M.R. (1991). Generalised linear models with random effects; a Gibbs Sampling approach. Journal American Statistical Society, 86,79-102