
References

- Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions*, Dover, New York.
- Accardi, L., Cabrera, J. and Watson, G. S. (1987). Some stationary Markov processes in discrete time for unit vectors, *Metron* **45**, 115–133.
- Aerts, M., Geys, H., Molenberghs, G. and Ryan, L. M. (2002). *Topics in Modelling of Clustered Data*, Chapman & Hall, London.
- Albert, P. S. (1999). Tutorial in biostatistics: Longitudinal data analysis (repeated measures) in clinical trials, *Statistics in Medicine* **18**, 1707–1732.
- Albert, P. S. and Waclawiw, M. A. (1998). A two-state markov chain for heterogeneous transitional data: a quasi-likelihood approach, *Statistics in Medicine* **17**, 1481–1493.
- Anderson, J. A. and Pemberton, J. D. (1985). The grouped continuous model for multivariate ordered categorical variables and covariate adjustment, *Biometrics* **41**, 875–885.
- Arnold, S. F. (1990). *Mathematical Statistics*, Prentice-Hall, New Jersey.
- Artes, R., Paula, G. A. and Ranvaud, R. (2000). Analysis of circular longitudinal data based on generalized estimating equations, *Australian & New Zealand Journal of Statistics* **42**, 347–358.
- Bahadur, R. R. (1961). A representation of the joint distribution of responses on n dichotomous items, *Studies in Item Analysis and Prediction, Stanford Mathematical Studies in the Social Sciences (H. Solomon, ed.)* **VI**, 158–168.
- Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*, Chapman & Hall, London.
- Barnhart, H. X. and Williamson, J. M. (1998). Goodness-of-fit tests for GEE modelling with binary responses, *Biometrics* **54**, 720–729.
- Beal, S. L. and Sheiner, L. B. (1988). Heteroscedastic nonlinear regression, *Technometrics* **30**, 327–338.
- Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge.

- Booth, J. G. and Hobert, J. P. (1999). Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm, *Journal of the Royal Statistical Society, Series B* **61**, 265–285.
- Borgioli, C., Martelli, M., Porri, F., D’Elia, A., Marchetti, G. M. and Scapini, F. (1999). Orientation in *talitrus saltator* (montagu): trends in intrapopulation variability related to environmental and intrinsic factors, *Journal of Experimental Marine Biology and Ecology* **238**, 29–47.
- Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- Breckling, J. (1989). *Analysis of Directional Time Series: Application to Wind Speed and Direction (Lecture Notes in Statistics 61)*, Springer-Verlag, Berlin.
- Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association* **88**, 9–25.
- Breslow, N. E. and Day, N. E. (1980). *Statistical Methods in Cancer Research*, IARC Scientific Publications No.32, Lyon.
- Breslow, N. E. and Lin, X. (1995). Bias correction in generalised linear mixed models with a single component of dispersion, *Biometrika* **82**, 81–91.
- Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd edn, Springer, New York.
- Brockwell, P. J. and Davis, R. A. (1996). *Introduction to Time Series and Forecasting*, Springer, New York.
- Brooks, S. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics* **7**, 434–455.
- Butler, S. M. and Louis, T. A. (1992). Random-effects models with non-parametric priors, *Statistics in Medicine* **11**, 1981–2000.
- Carey, V., Zeger, S. L. and Diggle, P. J. (1993). Modelling multivariate binary data with alternating logistic regression, *Biometrika* **80**, 517–526.
- Carlin, B. P. and Polson, N. G. (1992). Monte Carlo Bayesian methods for discrete regression models and categorical time series, *Bayesian Statistics* **4**, 577–586.
- Carlin, B. P., Polson, N. G. and Stoffer, D. S. (1992). A Monte Carlo approach to nonnormal and nonlinear state-space modeling, *Journal of the American Statistical Association* **87**, 493–500.
- Chaganty, N. R. (1997). An alternative approach to the analysis of longitudinal data via generalized estimating equations, *Journal of Statistical Planning and Inference* **63**, 39–54.
- Chaganty, N. R. and Joe, H. (2004). Efficiency of generalized estimating equations for binary responses, *Journal of the Royal Statistical Society, Series B* **66**, 851–860.
- Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions, *Journal of Econometrics* **34**, 305–334.
- Chan, J. S. K., Kuk, A. Y. C., Bell, J. and McGilchrist, C. (1998). The analysis of methadone clinic data using marginal and conditional logistic models

- with mixture or random effects, *Australian & New Zealand Journal of Statistics* **40**, 1–10.
- Chan, K. S. and Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving observations, *Journal of the American Statistical Association* **90**, 242–252.
- Chen, H. Y. and Little, R. J. A. (1995). A test of missing completely at random for generalised estimating equations with missing data, *Biometrika* **86**, 1–13.
- Coles, S. (1998). Inference for circular distributions and processes, *Statistics and Computing* **8**, 105–113.
- Cook, R. D. (1986). Assessment of local influence, *Journal of the Royal Statistical Society, Series B* **48**, 133–169.
- Cook, R. J. and Ng, E. T. M. (1997). A logistic-bivariate normal model for overdispersed two-state markov processes, *Biometrics* **53**, 358–364.
- Cook, R. J., Zeng, L. and Yi, G. Y. (2004). Marginal analysis of incomplete longitudinal binary data: a cautionary note on LOCF imputation, *Biometrics* **60**, 820–828.
- Cox, D. R. (1970). *The Analysis of Binary Data*, Chapman & Hall, London.
- Cox, D. R. (1981). Statistical analysis of time series, some recent developments, *Scandinavian Journal of Statistics* **8**, 93–115.
- Crowder, M. (1986). On consistency and inconsistency of estimating equations, *Econometric Theory* **3**, 305–330.
- Crowder, M. (1987). On linear and quadratic estimating function, *Biometrika* **74**, 591–597.
- Crowder, M. (1995). On the use of a working correlation matrix in using generalized linear models for repeated measures, *Biometrika* **82**, 407–410.
- Czado, C. and Song, P. X.-K. (2007). State space mixed models for longitudinal observations with binary and binomial responses, *Statistical Papers* **48**, to appear.
- Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measures*, Chapman & Hall, London.
- Davis, C. S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*, Springer, New York.
- Davis, P. J. and Rabinowitz, P. (1984). *Methods of Numerical Integration*, Academic Press, New York.
- de Jong, P. (1991). The diffuse Kalman filter, *Annals of Statistics* **2**, 1073–1083.
- de Jong, P. and Shephard, N. (1995). The simulation smoother for time series models, *Biometrika* **82**, 339–350.
- D’Elia, A., Borgioli, C. and Scapini, F. (2001). Orientation of sandhoppers under natural conditions in repeated trials: an analysis using longitudinal directional data, *Estuarine, Coastal and Shelf Science* **53**, 839–847.
- Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum-likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society, Series B* **39**, 1–38.

- Devroye, L. (1984). A simple algorithm for generating random variates with a log-concave density, *Computing* **33**, 247–257.
- Diggle, P. J. (1989). Testing for random dropouts in repeated measurement data, *Biometrics* **45**, 1255–1258.
- Diggle, P. J. (1990). *Time Series: An Biostatistical Introduction*, Clarendon Press, Oxford.
- Diggle, P. J. and Kenward, M. G. (1994). Informative dropout in longitudinal data analysis (with discussion), *Applied Statistics* **43**, 49–93.
- Diggle, P. J., Heagerty, P., Liang, K.-Y. and Zeger, S. L. (2002). *Analysis of Longitudinal Data*, 2nd edn, Oxford University Press, New York.
- Dragow, F. (1988). Polychoric and polyserial correlations, in L. Kotz and N. Johnson (eds), *Encyclopedia of Statistical Sciences*, Vol. 7, Wiley, New York, pp. 69–74.
- Durbin, J. (1960). Estimation of parameters in time-series regression models, *Journal of the Royal Statistical Society, Series B* **22**, 139–153.
- Durbin, J. and Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models, *Biometrika* **84**, 669–684.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, Chapman & Hall, London.
- Fahrmeir, L. (1992). Posterior mode estimation by extended Kalman filtering for multivariate dynamic generalized linear models, *Journal of the American Statistical Association* **87**, 501–509.
- Fahrmeir, L. and Kaufmann, H. (1987). Regression models for nonstationary categorical time series, *Journal of time Series Analysis* **8**, 147–160.
- Fahrmeir, L. and Tutz, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models*, Springer, Berlin.
- Fan, J. and Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications*, Chapman & Hall, London.
- Ferguson, T. S. (1958). A method of generating best asymptotically normal estimates with application to the estimation of bacterial densities, *Annals of Mathematical Statistics* **29**, 1046–1062.
- Fisher, N. I. (1993). *Statistical Analysis of Circular Data*, University Press, Cambridge.
- Fisher, N. I. and Lee, A. J. (1992). Regression models for an angular response, *Biometrics* **48**, 665–677.
- Fisher, N. I. and Lee, A. J. (1994). Time series analysis for circular data, *Journal of the Royal Statistical Society, Series B* **56**, 327–339.
- Fisher, R. A. (1935). The fiducial argument in statistical inference, *Annals of Eugenics* **6**, 391–398.
- Fitzmaurice, G. M. and Laird, N. M. (1995). Regression models for a bivariate discrete and continuous outcome with clustering, *Journal of the American Statistical Association* **90**, 845–852.

- Fitzmaurice, G. M. and Laird, N. M. (2000). Generalized linear mixture models for handling nonignorable dropouts in longitudinal studies, *Biostatistics* **1**, 141–156.
- Fitzmaurice, G. M., Laird, N. M. and Rotnitzky, A. (1993). Regression models for discrete longitudinal responses, *Statistical Science* **8**, 284–309.
- Fitzmaurice, G. M., Laird, N. M. and Ware, J. H. (2004). *Applied Longitudinal Analysis*, Wiley, New York.
- Fitzmaurice, G. M., Laird, N. M. and Zahner, E. P. (1996). Multivariate logistic models for incomplete binary responses, *Journal of the American Statistical Association* **91**, 99–108.
- Follmann, D. A. and Wu, M. C. (1995). An approximate generalized linear model with random-effects for informative missing data, *Biometrics* **51**, 151–168.
- Gelfand, A. E. and Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**, 398–409.
- Gelfand, A. E., Sahu, S. K. and Carlin, B. P. (1995). Efficient parameterizations for normal linear mixed models, *Biometrika* **82**, 479–488.
- Gelfand, A. E., Sahu, S. K. and Carlin, B. P. (1996). Efficient parametrizations for generalized linear mixed models, *Bayesian Statistics* **5**, 165–180.
- Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences (with discussion), *Statistical Science* **7**, 457–511.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.
- Genz, A. (1992). Numerical computation of multivariate normal probabilities, *Journal of Computational and Graphical Statistics* **1**, 141–149.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, in A. D. J.O. Berger, J.M. Bernardo and A. Smith (eds), *Bayesian Statistics 4*, Oxford University Press, Oxford, pp. 169–194.
- Geyer, C. J. and Thompson, E. A. (1992). Constrained Monte Carlo maximum likelihood for dependent data, *Journal of the Royal Statistical Society, Series B* **54**, 657–699.
- Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*, Chapman & Hall, New York.
- Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation, *Annals of Mathematical Statistics* **31**, 1208–1212.
- Godambe, V. P. (1976). conditional likelihood and unconditional optimum estimating equations, *Biometrika* **63**, 277–284.
- Godambe, V. P. (1991). *Estimating Functions*, Oxford University Press, Oxford.
- Godambe, V. P. and Thompson, M. E. (1974). Estimating equations in the presence of a nuisance parameter, *Annals of Statistics* **2**, 568–571.
- Godambe, V. P. and Thompson, M. E. (1978). Some aspects of the theory of estimating equations, *Journal of Statistical Planning and Inference* **2**, 95–104.

- Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models, *International Statistical Review* **55**, 245–259.
- Hall, A. R. (2005). *Generalized Method of Moments*, Oxford University Press, Oxford.
- Hansen, L. (1982). Large sample properties of generalized methods of moments estimators, *Econometrica* **50**, 1029–1055.
- Hardin, J. W. and Hilbe, J. M. (2003). *Generalized Estimating Equations*, Chapman & Hall, Boca Raton, FL.
- Harris, B. (1988). Tetrachoric correlation coefficient, in L. Kotz and N. Johnson (eds), *Encyclopedia of Statistical Sciences*, Vol. 9, Wiley, New York, pp. 223–225.
- Harvey, A. C. (1981). *Time Series Models*, Allan, Oxford.
- Harvey, A. C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems, *Journal of the American Statistical Association* **72**, 320–340.
- Heagerty, P. J. and Zeger, S. L. (1996). Marginal regression models for clustered ordinal measurements, *Journal of the American Statistical Association* **91**, 1024–1036.
- Heagerty, P. J. and Zeger, S. L. (1998). Lorelogram: a regression approach to exploring dependence in longitudinal categorical responses, *Journal of the American Statistical Association* **93**, 150–162.
- Hedeker, D. and Mermelstein, R. J. (2000). Analysis of longitudinal substance use outcomes using ordinal random-effects regression models, *Journal of Addiction* **3**, 381–394.
- Heidelberger, P. and Welch, P. (1983). Simulation run length control in the presence of an initial transient, *Operations Research* **31**, 1109–1144.
- Heyde, C. C. (1997). *Quasi-likelihood and Its Application*, Springer-Verlag, New York.
- Hogan, J. W. and Laird, N. M. (1997). Mixture models for the joint distribution of repeated measures and event times, *Statistics in Medicine* **16**, 239–258.
- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe, *Journal of the American Statistical Association* **47**, 663–685.
- Ibrahim, J. G., Chen, M.-H. and Lipsitz, S. R. (2001). Missing responses in generalized linear mixed models when the missing data mechanism is nonignorable, *Biometrika* **88**, 551–564.
- Ishwaran, H. and Takahara, G. (2002). Independent and identically distributed Monte Carlo algorithms for semiparametric linear mixed models, *Journal of the American Statistical Association* **97**, 1154–1166.
- Joe, H. (1995). Approximations to multivariate normal rectangle probabilities based on conditional expectations, *Journal of the American Statistical Association* **90**, 957–964.

- Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- Jones, B. and Kenward, M. G. (1989). *Design and Analysis of Cross-Over Trials*, Chapman & Hall, London.
- Jørgensen, B. (1987). Exponential dispersion models (with discussion), *Journal of the Royal Statistical Society, Series B* **49**, 127–162.
- Jørgensen, B. (1997). *The Theory of Dispersion Models*, Chapman & Hall, London.
- Jørgensen, B. and Knudsen, S. J. (2004). Parameter orthogonality and bias adjustment for estimating functions, *Scandinavian Journal of Statistics* **31**, 93–114.
- Jørgensen, B. and Song, P. X.-K. (1998a). Stationary time series models with exponential dispersion model margins, *Journal of Applied Probability* **35**, 78–92.
- Jørgensen, B. and Song, P. X.-K. (1998b). Stationary state space models for longitudinal data, *Research Report #9*, Department of Statistics, University of Southern Denmark.
- Jørgensen, B., Lundbye-Christensen, S., Song, P. X.-K. and Sun, L. (1996a). State-space models for multivariate longitudinal data of mixed types, *Canadian Journal of Statistics* **24**, 385–402.
- Jørgensen, B., Lundbye-Christensen, S., Song, P. X.-K. and Sun, L. (1996b). A longitudinal study of emergency room visits and air pollution for Prince George, British Columbia, *Statistics in Medicine* **15**, 823–836.
- Jørgensen, B., Lundbye-Christensen, S., Song, P. X.-K. and Sun, L. (1999). A state-space models for multivariate longitudinal count data, *Biometrika* **86**, 169–181.
- Jørgensen, B., Martinez, J. R. and Tsao, M. (1994). Asymptotic behaviour of the variance function, *Scandinavian Journal of Statistics* **21**, 223–243.
- Kalbfleisch, J. D. and Lawless, J. F. (1985). The analysis of panel data under a markov assumption, *Journal of the American Statistical Association* **80**, 863–873.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *Journal of Basic Engineering* **82**, 34–45.
- Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction theory, *Journal of Basic Engineering* **83**, 95–108.
- Karim, M. R. and Zeger, L. R. (1988). GEE: a SAS macro for longitudinal data analysis, *Technical Report #674*, Department of Biostatistics, The Johns Hopkins University, Baltimore.
- Kass, R. and Raftery, A. (1995). Bayes factors and model uncertainty, *Journal of the American Statistical Association* **90**, 773–795.
- Kendall, M. G. and Gibbons, J. D. (1990). *Rank Correlation Methods*, Oxford University Press, New York.
- Kimball, B. F. (1946). Sufficient statistical estimation functions for the parameters of the distribution of maximum values, *Annals of Mathematical Statistics* **17**, 299–309.
- Kitagawa, G. (1987). Non-Gaussian state-space modelling of non stationary time series (with discussion), *Journal of the American Statistical Association* **82**, 1032–1063.

- Klein, J. P. and Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*, Springer, New York.
- Knight, K., Leroux, B. G., Millar, J. and Petkau, A. (1989). Air pollution and human health: A study based on emergency room visits data from Prince George, British Columbia, *Technical Report #136*, Department of Statistics, University of British Columbia.
- Kosorok, M. R. and Chao, W. H. (1996). The analysis of longitudinal ordinal response data in continuous time, *Journal of the American Statistical Association* **91**, 807–817.
- Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data, *Biometrics* **38**, 963–974.
- Lambert, M. L., Lamble, S. J. and Girard, R. W. (1987). *1986 Annual Air Quality Report for Prince George*, Waste Management Branch, Ministry of Environment & Parks of British Columbia, Prince George, British Columbia.
- Lange, K. (1999). *Numerical Analysis for Statisticians*, Springer-Verlag, New York.
- Lee, M. J. (1996). *Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models*, Springer, New York.
- Lee, Y. and Nelder, J. A. (1996). Hierarchical generalized linear models (with discussion), *Journal of the Royal Statistical Society, Series B* **58**, 619–678.
- Lee, Y. and Nelder, J. A. (2004). Conditional and marginal models: Another view, *Statistical Science* **19**, 219–238.
- Lewis, P. A. W., McKenzie, E. and Hugus, D. K. (1989). Gamma processes, *Communication in Statistics—Stochastic Models* **5(1)**, 1–30.
- Liang, K.-Y. and Self, S. G. (1996). On the asymptotic behaviour of the pseudo-likelihood ratio test statistic, *Journal of the Royal Statistical Society, Series B* **58**, 785–796.
- Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models, *Biometrika* **73**, 13–22.
- Liao, J. G. and Qaqish, B. F. (2005). Discussion of “maximization by parts in likelihood inference” by Song, Fan, and Kalbfleisch, *Journal of the American Statistical Association* **100**, 1160–1161.
- Lin, X. and Breslow, N. E. (1996). Bias correction in generalized linear mixed models with multiple components of dispersion, *Journal of the American Statistical Association* **91**, 1007–1016.
- Lindsay, B. G. (1988). Composite likelihood methods, *Contemporary Mathematics* **80**, 221–240.
- Lindsay, B. G. and Qu, A. (2003). Inference functions and quadratic score tests, *Statistical Science* **18**, 394–410.
- Lindsey, J. K. (1999). *Models for Repeated Measurements*, 2nd edn, Oxford University Press, Oxford.
- Lindstrom, M. J. and Bates, D. M. (1988). Newton-Raphson and EM-algorithms for linear mixed-effects models for repeated-measures data, *Journal of the American Statistical Association* **83**, 1014–1022.

- Lindstrom, M. J. and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data, *Biometrics* **46**, 673–687.
- Lipsitz, S. R. and Ibrahim, J. G. (2000). Estimation with correlated censored survival data with missing covariates, *Biostatistics* **1**, 315–327.
- Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data, *Journal of the American Statistical Association* **88**, 125–134.
- Little, R. J. A. (1994). A class of pattern-mixture models for normal incomplete data, *Biometrika* **81**, 471–483.
- Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis With Missing Data*, 2nd edn, Wiley, New York.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm, *Journal of the Royal Statistical Society, Series B* **44**, 226–233.
- Mardia, K. V. (1972). *Statistics of Directional Data*, Academic Press, London.
- McCullagh, P. (1983). Quasi-likelihood functions, *Annals of Statistics* **11**, 59–67.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd edn, Chapman & Hall, London.
- McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models, *Journal of the American Statistical Association* **92**, 162–170.
- McCulloch, C. E. and Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*, Wiley & Sons, New York.
- McGilchrist, C. A. (1994). Estimation in generalized mixed models, *Journal of the Royal Statistical Society, Series B* **56**, 61–69.
- McLeish, D. L. and Small, C. G. (1988). *The Theory and Applications of statistical Inference Functions*, Lecture Notes in Statistics 44, Springer-Verlag, New York.
- Meyers, S. M., Ambler, J. S., Tan, M., Werner, J. C. and Huang, S. S. (1992). Variation of perfluoropropane disappearance after vitrectomy, *Retina* **12**, 359–363.
- Minini, P. and Chavance, M. (2004). Sensitivity analysis of longitudinal binary data with non-monotone missing values, *Biostatistics* **5**, 531–544.
- Molenberghs, G., Kenward, M. G. and Lesaffre, E. (1997). The analysis of longitudinal ordinal data with nonrandom drop-out, *Biometrika* **84**, 33–44.
- Muenz, L. R. and Rubinstein, L. V. (1985). Markov models for covariate dependence of binary sequence, *Biometrics* **43**, 863–871.
- Neal, R. (1998). Suppressing random walks in Markov chain Monte Carlo using ordered over-relaxation, in M. Jordan (ed.), *Learning in Graphical Models*, Kluwer Academic Publishers, Dordrecht, pp. 205–230.
- Neuhaus, J. M. and Kalbfleisch, J. D. (1998). Between- and within-cluster covariate effects in the analysis of clustered data, *Biometrics* **54**, 638–645.
- Neuhaus, J. M., Hauck, W. W. and Kalbfleisch, J. D. (1992). The effects of mixture distribution misspecification when fitting mixed-effects logistic models, *Biometrika* **1992**, 755–762.
- Neuhaus, J. M., Kalbfleisch, J. D. and Hauck, W. W. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data, *International Statistics Review* **59**, 25–35.

- Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing, *Handbook of Econometrics* **4**.
- Neyman, J. (1949). Contribution to the theory of χ^2 test, in J. Neyman (ed.), *Berkeley Symposium in Mathematical Statistics and Probability*, University of California Press, Berkeley, CA, pp. 239–273.
- Olkin, I. and Pratt, J. W. (1958). A multivariate Tchebycheff inequality, *Annals of Mathematical Statistics* **29**, 226–234.
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient, *Psychometrika* **44**, 443–460.
- Paik, M. C. (1992). Parametric variance function estimation for nonnormal repeated measurement data, *Biometrics* **48**, 19–30.
- Paik, M. C. (1997). The generalized estimating equation approach when data are not missing completely at random, *Journal of the American Statistical Association* **92**, 1320–1329.
- Pan, W. (2001). Model selection in estimating equations, *Biometrics* **57**, 529–534.
- Pan, W. (2002). Goodness-of-fit tests for GEE with correlated binary data, *Scandinavian Journal of Statistics* **29**, 101–110.
- Penrose, K., Nelson, A. and Fisher, A. (1985). Generalized body composition prediction equation for men using simple measurement techniques (abstract), *Medicine and Science in Sports and Exercise* **17**, 189.
- Petkau, A. J. and White, R. A. (2003). Statistical approaches to assessing the effects of neutralizing antibodies: IFNbeta-1b in the pivotal trial of relapsing-remitting multiple sclerosis, *Neurology* **61(9 Suppl 5)**, 35–37.
- Petkau, A. J., White, R. A., G. C. Ebers, A. T. R., Sibley, W. A., Lublin, F. D. and Paty, D. W. (2004). Longitudinal analysis of the effects of neutralizing antibodies on interferon beta-1b in relapsing-remitting multiple sclerosis, *Multiple Sclerosis* **10**, 126–138.
- Pinheiro, J. C., Liu, C. and Wu, Y. N. (2001). Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution, *Journal of Computational and Graphical Statistics* **10**, 249–276.
- Preisser, J. S. and Qaqish, B. F. (1996). Deletion diagnostics for generalized estimating equations, *Biometrika* **83**, 551–562.
- Prentice, R. L. (1988). Correlated binary regression with covariates specific to each binary observation, *Biometrics* **44**, 1033–1048.
- Prentice, R. L. and Zhao, L. P. (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses, *Biometrics* **47**, 825–839.
- Qiu, Z., Song, P. X.-K. and Tan, M. (2002). Bayesian hierarchical analysis of multi-level repeated ordinal data using winbugs, *Journal of Biopharmaceutical Statistics* **12**, 121–135.
- Qu, A. and Song, P. X.-K. (2002). Testing ignorable missingness in estimating equation, *Biometrika* **89**, 841–850.

- Qu, A. and Song, P. X.-K. (2004). Assessing robustness of generalised estimating equations and quadratic inference functions, *Biometrika* **91**, 447–459.
- Qu, A., Lindsay, B. G. and Li, B. (2000). Improving generalized estimating equations using quadratic inference functions, *Biometrics* **87**, 823–836.
- Qu, Y. and Tan, M. (1998). Analysis of clustered ordinal data with subclusters via a Bayesian hierarchical model, *Communications in Statistics: Theory and Methods* **27**, 1461–1475.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edn, Wiley, New York.
- Raudenbush, S. W., Yang, M.-L. and Yosef, M. (2000). Maximum likelihood for generalized linear models with nested random effects via high-order, multivariate Laplace approximation, *Journal of Computational and Graphical Statistics* **9**, 141–157.
- Rivest, L.-P. (1997). A decentred predictor for circular-circular regression, *Biometrika* **84**, 717–726.
- Robert, C. P. (1995). Simulation of truncated normal variables, *Statistics and Computing* **5**, 121–125.
- Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed, *Journal of the American Statistical Association* **89**, 846–866.
- Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes under the presence of missing data, *Journal of the American Statistical Association* **90**, 106–121.
- Rotnitzky, A. and Jewell, N. (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data, *Biometrika* **90**, 485–497.
- Rubin, D. B. (1976). Inference and missing data, *Biometrika* **63**, 581–592.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*, Wiley, New York.
- Rubin, D. B., Stern, H. S. and Vehovar, V. (1994). Handling “don’t know” survey responses: the case of the slovenian plebiscite, *Journal of the American Statistical Association* **90**, 822–828.
- Ruppert, D. (2005). Discussion of “maximization by parts in likelihood inference” by Song, Fan, and Kalbfleisch, *Journal of the American Statistical Association* **100**, 1161–1163.
- Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*, Chapman & Hall, London.
- Schall, R. (1999). Estimation in generalized linear models with random effects, *Biometrika* **78**, 719–727.
- Shih, W. J. and Quan, H. (1997). Testing for treatment differences with dropouts present in clinical trials—a composite approach, *Statistics in Medicine* **16**, 1225–1239.

- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de L'Institut de Statistiques de l'Université de Paris* **8**, 229–231.
- Sommer, A., Katz, J. and Tarwotjo, I. (1984). Increased risk of respiratory infection and diarrhea in children with pre-existing mild vitamin A deficiency, *American Journal of Clinical Nutrition* **40**, 1090–1095.
- Song, P. X.-K. (2000a). Multivariate dispersion models generated from Gaussian copula, *Scandinavian Journal of Statistics* **27**, 305–320.
- Song, P. X.-K. (2000b). Monte Carlo Kalman filter and smoothing for multivariate discrete state space models, *Canadian Journal of Statistics* **28**, 641–652.
- Song, P. X.-K. and Tan, M. (2000). Marginal models for longitudinal continuous proportional data, *Biometrics* **56**, 496–502.
- Song, P. X.-K., Fan, Y. and Kalbfleisch, J. D. (2005). Maximization by parts in likelihood inference (with discussion), *Journal of the American Statistical Association* **100**, 1145–1167.
- Song, P. X.-K., Qiu, Z. and Tan, M. (2004). Modelling heterogeneous dispersion in marginal simplex models for longitudinal continuous proportional data, *Biometrical Journal* **46**, 540–553.
- Song, P. X.-K., Zhang, P. and Qu, A. (2007). Maximum likelihood inference in robust linear mixed-effects models using multivariate t distributions, *Statistica Sinica* **17**, to appear.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion), *Journal of the Royal Statistical Society, Series B* **64**, 583–639.
- Stiratelli, R., Laird, N. M. and Ware, J. H. (1984). Random-effects models for serial observations with binary response, *Biometrics* **40**, 961–971.
- Stoffer, D. S., Scher, M. S., Richardson, G. A., Day, N. L. and Coble, P. A. (1988). A Walsh-Fourier analysis of the effects of moderate maternal alcohol consumption on neonatal sleep-state cycling, *Journal of the American Statistical Association* **83**, 954–963.
- Sun, D., Tsutakawa, R. K. and He, Z. (2001). Propriety of posteriors with improper priors in hierarchical linear mixed models, *Statistica Sinica* **11**, 77–95.
- Sun, J. and Song, P. X.-K. (2001). Statistical analysis of repeated measurements with informative censoring times, *Statistics in Medicine* **20**, 63–73.
- Sutradhar, B. C. and Dass, K. (1999). On the efficiency of regression estimators in generalized linear models for longitudinal data, *Biometrika* **86**, 459–465.
- Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association* **82**, 528–540.
- Thall, P. F. and Vail, S. C. (1990). Some covariance models for longitudinal count data with overdispersion, *Biometrics* **46**, 657–671.
- van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*, Springer-Verlag, New York.

- Verbeke, G. and Lesaffre, E. (1997). The effect of misspecifying the random effects distribution in linear mixed models for longitudinal data, *Journal of Computational Statistics and Data Analysis* **23**, 541–556.
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*, Springer-Verlag, New York.
- Vonesh, E. F. and Carter, R. L. (1987). Efficient inference for random-coefficient growth curve models with unbalanced data, *Biometrics* **43**, 617–628.
- Wang, Y.-G. and Carey, V. J. (2003). Working correlation structure misspecification, estimation and covariance design: Implications for the GEE performance, *Biometrika* **90**, 29–41.
- Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss-Newton method, *Biometrika* **61**, 439–447.
- Wei, G. C. G. and Tanner, M. A. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithm, *Journal of the American Statistical Association* **85**, 699–704.
- Weiss, R. E. and Lazaro, C. G. (1992). Residual plots for repeated measures, *Statistics in Medicine* **11**, 115–124.
- West, M. and Harrison, P. J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd edn, Springer-Verlag, New York.
- Wu, H. and Zhang, J.-T. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches*, Wiley, New York.
- Wu, M. C. and Carroll, R. J. (1988). Estimation and comparison of changes in the presence of informative right censoring: conditional linear model, *Biometrics* **44**, 175–188.
- Xing, B. (2004). *Best Quadrature Formulas, Mixture of Normal Approximation and State Space Models*, PhD thesis, Department of Mathematics and Statistics, York University, Toronto, Ontario.
- Yi, G. Y. and Cook, R. J. (2002). Marginal methods for incomplete longitudinal data arising in clusters, *Journal of the American Statistical Association* **97**, 1071–1080.
- Zeger, S. L. (1988). A regression model for time series of counts, *Biometrika* **75**, 621–629.
- Zeger, S. L. and Karim, R. (1991). Generalized linear models with random effects: a Gibbs sampling approach, *Journal of the American Statistical Association* **86**, 79–86.
- Zeger, S. L. and Qaqish, B. (1988). Markov regression models for time series: A quasi-likelihood approach, *Biometrics* **44**, 1019–1031.
- Zeger, S. L., Liang, K.-Y. and Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach, *Biometrics* **44**, 1049–1060.
- Zhang, D. and Davidian, M. (2001). Linear mixed models with flexible distributions of random effects for longitudinal data, *Biometrics* **57**, 795–802.
- Zhang, P. (2006). *Contributions to Mixed-Effects Models for Longitudinal Data*, PhD thesis, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario.

- Zhao, L. P. and Prentice, R. L. (1990). Correlated binary regression using a generalized quadratic model, *Biometrika* **77**, 642–648.
- Ziegler, A., Kastner, C. and Blettner, M. (1998). The generalized estimating equations: An annotated bibliography, *Biometrical Journal* **40**, 115–139.
- Zorn, C. J. W. (2001). Generalized estimating equation models for correlated data: A review with applications, *American Journal of Political Science* **45**, 470–490.

Index

- algorithm
 - EM, *see* EM algorithm
 - Fisher-scoring, 52, 187
 - Gauss-Newton, 137, 146, 170
 - iterative weighted least squares, 181
 - maximization by parts, 137, 142–146
 - Monte Carlo EM, 180
 - Newton-scoring, 106, 281
- angular data, 10
- AR(1) process, 239
- association, 75
 - Kendall's τ , 77, 129
 - normal scoring ν , 77, 130
 - odds ratio, 76
 - polychoric, 133
 - Spearman's ρ , 77, 129
 - tetrachoric, 133
- association matrix, *see* correlation structure, 10
- association model, 75
 - conditional, 75, 80, 157
 - joint, 75
 - quasi-likelihood, 75
- asymptotic relative efficiency, 148
- augmented likelihood, 265
- autocorrelation, 5
- available-case analysis, 299, 307

- Bahadur's representation, 122, 124
- Bayesian inference, 195
 - formulation, 195
 - posterior, 196
 - priors, 196, 248

 - summary statistics, 198
 - best linear unbiased predictor, 182, 217
 - BLUP, *see* best linear unbiased predictor
- Cholesky decomposition, 187
- comb structure, 222, 229
- complete-case analysis, 298, 307
- conditional residual, 282
- convolution, 34
- correlated data, 1
 - clustered data, 1
 - longitudinal data, 1
 - mixed type, 153
 - multi-level data, 2
 - repeated measurements, 2, 13
 - spatial data, 2
 - vector data, 2, 121
- correlation structure, *see* association
 - 1-dependence, 141
 - AR-1, 78, 97, 104, 141
 - exponential, 97
 - independence, 78
 - interchangeability, 78, 97, 104, 140, 162
 - m-dependence, 79
 - unstructured, 78, 140
- cross-over trial, 4, 150
- cross-sectional data,, 5
- Crowder optimality, 65
 - multi-dimensional, 68

- data example, 6
 - air pollution study, 16, 283–289

- body fat index, 49
- burn injury, 153
- children hospital visit, 152
- children's health studies, 6
- cross-over trial, 150
- epileptic seizures, 7, 112–115
- infant sleep, 255–259
- multiple sclerosis trial, 13, 110–112, 171–174, 203–206
- retinal surgery, 9, 116–120, 188–190
- sandhopper orientation, 10, 190–192
- schizophrenia trial, 11, 302, 304, 325–329
- shift-mean time series, 232
- TEC trial, 13, 207–211
- Tokyo rainfall, 15, 246, 252–255
- US polio incidences, 14, 272–274
- deviance information criterion, 249
- deviance score, 38
- dispersion model, 20, 26
 - asymptotic normality, 29, 186
 - bivariate exponential model, 143
 - maximum likelihood estimation, 37
 - multivariate, 128
 - binary, 132
 - continuous, 128
 - discrete, 128
 - gamma, 136
 - mixed type, 128
 - Poisson, 133
 - property, 28
 - saddlepoint approximation, 29
- duality transformation, 33
- EM algorithm, 311–317
 - linear regression, 316–317
 - Louis formula, 315
- estimating equation
 - stable, 55
- estimating function, *see* inference function
- exponential dispersion model, 24, 30
 - additive, 33
- filter density, 229
- Fisher information, 62
- Fréchet bound, 77, 127, 133
- gamma process, 262, 263
- Gaussian copula, 122
 - density, 126
 - parametric copula, 125
- GEE, 92, 261
 - ED, 93
 - GEE2, 95
 - constant dispersion, 96
 - varying dispersion, 100
 - inverse probability weighting, 318
 - parameter-driven, 264–265
 - simplex, 94
 - von Mises, 94
- generalized linear mixed effects model, 163
 - marginal interpretation, 162, 174
- generalized method of moment, 61, 68
- generalized state space model, 228
 - nonstationary, 263–264
 - stationary, 262–263
- Gibbs sampler, 181, 195–198, 267
- Godambe inequality, 64
 - multi-dimensional, 67
- Godambe information, 62, 271
 - sandwich covariance estimator, 92
 - sensitivity, 62
 - sensitivity matrix, 66
 - variability, 62
 - variability matrix, 66
- hierarchical model, 207
 - prior, 208
 - Wishart distribution, 209
 - proportional odds model, 207
- inference function, 20, 55
 - additive, 59
 - Crowder class, 65
 - equivalent, 59
 - generalized method of moments, 68
 - insensitivity, 62, 92, 186, 269, 281
 - kernel, 59
 - multi-dimensional, 65
 - normalized, 63, 67
 - optimal, 55, 61
 - regular, 61, 65
 - unbiased, 59
- interim model, 241
- inverse gamma, 248

- jumps, 232
- Kalman estimating equation, 228
 - nonstationary, 275–283
 - stationary, 267–272
- Kalman filter, 223, 270, 279
- Kalman smoother, 224, 270, 279
- Laplace method, 183
- linear mixed effects model, 161
- linear state space model, 231
 - Gaussian, 231
- listwise deletion, 298
- local influence, 293
- log-linear model representation, *see*
 - Bahadur’s representation
 - quadratic exponential model, 123
- lognormal model, 262
- long-term effects, 263
- longitudinal plot, 9
- lorelogram, 80, 112
- marginal GLM, 88
- Markov chain Monte Carlo, 195, 247
 - burn-in, 197
 - convergence, 197
 - convergence diagnostics, 198–201
 - autocorrelation plot, 199
 - Brooks-Gelman-Rubin R statistic, 199
 - Geweke test, 199–200
 - Heidelberger and Welch’s test, 200
 - trace plot, 199
 - enhancing burn-in, 201–202
 - hierarchical centering, 201
 - sweeping method, 201
 - Gibbs sampler, *see* Gibbs sampler
 - model selection, 202
 - deviance information criterion, 202
 - over-relaxation, 198
 - thinning, 197
- maximization by parts, 237
- maximum likelihood
 - conditional approach, 176–178
 - EM algorithm, 178–182
 - numerical integration, 167–174
 - PQL and REML, 182–192
 - REML, 182
 - simulated, 174–176
- mean square error, 217, 271
- mean value mapping, 31
- missing data, 6
 - dropout, 11
 - marginal model, 320
 - transition model, 319
 - ignorable, 303
 - informative, 6
 - MAR, 6, 297
 - MCAR, 6, 109, 297
 - NMAR, 297
 - pattern mixture model, 321
 - selection model, 320
 - shared model, 321
 - nonignorable, 303
 - score-type test, 303
- missing data pattern, 293–296
 - arbitrary, 295
 - complete, 294
 - monotonic, 295
 - uniform, 294
 - univariate, 294
- model diagnosis, 281–283
- model selection, 103
 - Akaike information criterion, 106
 - Bayes information criterion, 106
 - deviance information criterion, *see*
 - deviance information criterion
- Monte Carlo EM algorithm, 265–267
 - E-step, 266
 - M-step, 266
- Monte Carlo Kalman filter, 243
- Monte Carlo Kalman smoother, 242
- Monte Carlo MLE, 235
- multiple imputation, 307–311
 - between-imputation variance, 309
 - imputation variation, 308
 - sampling variation, 308
 - within-imputation variance, 309
- multiplicative AR(1) process, 262
- normal-score transformation, 77
- over-identification, 59
- overdispersion, 55, 56, 159
- parameter-driven model, 261–262
- penalized quasi-likelihood, 182, 185
- Poisson-gamma model, 275

- quadratic inference function, 103
 - robust, 103, 112
 - test for MCAR, 302
- quadrature method
 - Gauss-Hermite, 167
 - Hermite orthogonal polynomials, 168
- quasi-likelihood, 58
 - penalized, *see* penalized quasi-likelihood
- quasi-score, 57

- random effects, 82, 159
- random walk, 239
- residual, 35, 96
 - deviance, 35
 - model diagnosis, 101
 - modified score, 91
 - Pearson, 35, 96
- residual analysis, 79, 281

- saddlepoint, 183
- serial correlation, *see* autocorrelation
- shift-mean model, 232
- short-term effects, 263
- signal-to-noise ratio, 248
- simplex distribution, 10, 30, 42
- simulation smoother, 237
- smoothed best quadrature formula, 234
- smoother density, 230
- software, 106
 - R, 53
 - SAS, 107
 - macro qif, 108
 - proc genmod, 107
 - proc glimmix, 183, 194
 - proc mi, 308
 - proc mianalyze, 310
 - proc mixed, 166, 185, 192
 - proc nlmixed, 167, 171, 193
 - WinBUGS, 195, 212
 - BOA, 201
 - CODA, 201
 - Doodle, 203
- state space mixed model, 247
- stochastic representation, 134
- subject-specific effects, *see* random effects
- surrogate response, 181

- third order approximation, 36
 - Barndorff-Nielsen formula, 36
 - Lugannani-Rice formula, 36
- Tweedie class, 34, 276

- U statistic, 323
- unit deviance, 26
 - regular, 26
- unit variance function, 26
 - power, 36

- variable mesh, 234
- variogram, 79
- vector GLM, 121
 - reproducible, 122
- von Mises distribution, 50

- Wald test, 95
- Wilcoxon-type test, 323
- working correlation, 91

- Yule-Walker equation, 265