References

Abramowitz, M., Stegun, I.A., Danos, M. and Rafelski, J. (eds.) (1984). *Pocketbook of Mathematical Functions. Abridged Edition of Handbook of Mathematical Functions.* Verlag Harri Deutsch, Thun and Frankfurt am Main.

Aitchison, J. (1982). The statistical analysis of compositional data. *J. Roy. Statist. Soc.* B **44**, 139–177.

Albert, P.S. (1991). A two-state Markov mixture model for a time series of epileptic seizure counts. *Biometrics* **47**, 1371–1381.

Altman, R. MacKay (2007). Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting. *J. Amer. Statist. Assoc.* **102**, 201–210.

Altman, R. MacKay and Petkau, J.A. (2005). Application of hidden Markov models to multiple sclerosis lesion count data. *Statist. Med.* **24**, 2335–2344.

Aston, J.A.D. and Martin, D.E.K. (2007). Distributions associated with general runs and patterns in hidden Markov models. *Ann. Appl. Statist.* **1**, 585–611.

Azzalini, A. and Bowman, A.W. (1990). A look at some data on the Old Faithful geyser. *Appl. Statist.* **39**, 357–365.

Barton Browne, L. (1993). Physiologically induced changes in resource-oriented behaviour. *Ann. Rev. Entomology* **38**, 1–25.

Baum, L.E. (1972). An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes. In *Proc. Third Symposium on Inequalities*, O. Shisha (ed.), 1–8. Academic Press, New York.

Baum, L.E., Petrie, T., Soules, G. and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.* **41**, 164–171.

Bellman, R. (1960). *Introduction to Matrix Analysis.* McGraw-Hill, New York.

Berchtold, A. (1999). The double chain Markov model. *Commun. Stat. Theory Meth.* **28**, 2569–2589.

Berchtold, A. (2001). Estimation in the mixture transition distribution model. *J. Time Series Anal.* **22**, 379–397.

Berchtold, A. and Raftery, A.E. (2002). The mixture transition distribution model for high-order Markov chains and non-Gaussian time series. *Statist. Sci.* **17**, 328–356.

Bisgaard, S. and Travis, L.E. (1991). Existence and uniqueness of the solution of the likelihood equations for binary Markov chains. *Statist. Prob. Letters* **12**, 29–35.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning.* Springer, New York.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis, Forecasting and Control*, third edition. Prentice Hall, Englewood Cliffs, NJ.

Boys, R.J. and Henderson, D.A. (2004). A Bayesian approach to DNA sequence segmentation (with discussion). *Biometrics* **60**, 573–588.

Brockwell, A.E. (2007). Universal residuals: A multivariate transformation. *Statist. Prob. Letters* **77**, 1473–1478.

Bulla, J. and Berzel, A. (2008). Computational issues in parameter estimation for stationary hidden Markov models. *Computat. Statist.* **23**, 1–18.

Bulla, J. and Bulla, I. (2007). Stylized facts of financial time series and hidden semi-Markov models. *Computat. Statist. & Data Analysis* **51**, 2192–2209.

Calvet, L. and Fisher, A.J. (2001). Forecasting multifractal volatility. *J. Econometrics* **105**, 27–58.

Calvet L. and Fisher, A.J. (2004). How to forecast long-run volatility: regime switching and the estimation of multifractal processes. *J. Financial Econometrics* **2**, 49–83.

Cappé, O., Moulines, E. and Rydén, T. (2005). *Inference in Hidden Markov Models.* Springer, New York.

Celeux, G., Hurn, M. and Robert, C.P. (2000). Computational and inferential difficulties with mixture posterior distributions. *J. Amer. Statist. Assoc.* **95**, 957–970.

Chib, S. (1996). Calculating posterior distributions and modal estimates in Markov mixture models. *J. Econometrics* **75**, 79–97.

Chopin, N. (2007). Inference and model choice for sequentially ordered hidden Markov models. *J. Roy. Statist. Soc.* B **69**, 269–284.

Congdon, P. (2006). Bayesian model choice based on Monte Carlo estimates of posterior model probabilities. *Computat. Statist. & Data Analysis* **50**, 346–357.

Cook, R.D. and Weisberg, S. (1982). *Residuals and Influence in Regression.* Chapman & Hall, London.

Cosslett, S.R. and Lee, L.-F. (1985). Serial correlation in latent discrete variable models. *J. Econometrics* **27**, 79–97.

Cox, D.R. (1981). Statistical analysis of time series: some recent developments. *Scand. J. Statist.* **8**, 93–115.

Cox, D.R. (1990). Role of models in statistical analysis. *Statist. Sci.* **5**, 169–174.

Cox, D.R. and Snell, E.J. (1968). A general definition of residuals (with discussion). *J. Roy. Statist. Soc.* B **30**, 248–275.

Davison, A.C. (2003). *Statistical Models.* Cambridge University Press, Cambridge.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc.* B **39**, 1–38.

Dewsbury, D.A. (1992). On the problems studied in ethology, comparative psychology, and animal behaviour. *Ethology* **92**, 89–107.

Diggle, P.J. (1993). Contribution to the discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo methods. *J. Roy. Statist. Soc.* B **55**, 67–68.

Draper, D. (2007). Contribution to the discussion of Raftery *et al.* (2007) (pp. 36–37).

Dunn, P.K. and Smyth, G.K. (1996). Randomized quantile residuals. *J. Comp. Graphical Statist.* **5**, 236–244.

Durbin, R., Eddy, S.R., Krogh, A. and Mitchison, G. (1998). *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids.* Cambridge University Press, Cambridge.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, New York.

Ephraim, Y. and Merhav, N. (2002). Hidden Markov processes. *IEEE Trans. Inform. Th.* **48**, 1518–1569.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Volume 1*, third edition. Wiley, New York.

Fisher, N.I. (1993). *The Analysis of Circular Data.* Cambridge University Press, Cambridge.

Fisher, N.I. and Lee, A.J. (1983). A correlation coefficient for circular data. *Biometrika* **70**, 327–332.

Fisher, N.I. and Lee, A.J. (1994). Time series analysis of circular data. *J. Roy. Statist. Soc.* B **56**, 327–339.

Forney, G.D. (1973). The Viterbi algorithm. *Proc. IEEE* **61**, 268–278.

Franke, J. and Seligmann, T. (1993). Conditional maximum-likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In *Developments in Time Series Analysis*, T. Subba Rao (ed.), 310–330. Chapman & Hall, London.

Fredkin, D.R. and Rice, J.A. (1992). Bayesian restoration of single-channel patch clamp recordings. *Biometrics* **48**, 427–448.

Fridman, M. and Harris, L. (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *J. Bus. Econ. Statist.* **16**, 284–291.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models.* Springer, New York.

Gill, P.E., Murray, W., Saunders, M.A. and Wright, M.H. (1986). User's Guide for NPSOL: a Fortran package for nonlinear programming. Report SOL 86-2, Department of Operations Research, Stanford University.

Gill, P.E., Murray, W. and Wright, M.H. (1981). *Practical Optimization.* Academic Press, London.

Gould, S.J. (1997). *The Mismeasure of Man*, revised and expanded edition. Penguin Books, London.

Granger, C.W.J. (1982). Acronyms in time series analysis (ATSA). *J. Time Series Anal.* **3**, 103–107.

Green, P.J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711–732.

Grimmett, G.R. and Stirzaker, D.R. (2001). *Probability and Random Processes*, third edition. Oxford University Press, Oxford.

Guttorp, P. (1995). *Stochastic Modeling of Scientific Data.* Chapman & Hall, London.

Haines, L.M., Munoz, W.P. and van Gelderen, C.J. (1989). ARIMA modelling of birth data. *J. Appl. Statist.* **16**, 55–67.

Haney, D.J. (1993). Methods for analyzing discrete-time, finite state Markov chains. Ph.D. dissertation, Department of Statistics, Stanford University.

Harte, D. (2008). **R** package 'HiddenMarkov', version 1.2-5. URL http://www.statsresearch.co.nz, accessed 27 July 2008.

Hasselblad, V. (1969). Estimation of finite mixtures of distributions from the exponential family. *J. Amer. Statist. Assoc.* **64**, 1459–1471.

Hastie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models.* Chapman & Hall, London.

Hastie, T., Tibshirani, R.J. and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction.* Springer, New York.

Holzmann, H., Munk, A., Suster, M.L. and Zucchini, W. (2006). Hidden Markov models for circular and linear-circular time series. *Environ. Ecol. Stat.* **13**, 325–347.

Hopkins, A., Davies, P. and Dobson, C. (1985). Mathematical models of patterns of seizures: their use in the evaluation of drugs. *Arch. Neurol.* **42**, 463–467.

Hughes, J.P. (1993). A class of stochastic models for relating synoptic atmospheric patterns to local hydrologic phenomena. Ph.D. dissertation, University of Washington.

Ihaka, R. and Gentleman, R. (1996). R: a language for data analysis and graphics. *J. Comp. Graphical Statist.* **5**, 299–314.

Jackson, C.H., Sharples, L.D., Thompson, S.G., Duffy, S.W. and Couto, E. (2003). Multistate Markov models for disease progression with classification error. *The Statistician* **52**, 193–209.

Jacquier, E., Polson, N.G. and Rossi, P.E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *J. Econometrics* **122**, 185–212.

Jammalamadaka, S.R. and Sarma, Y.R. (1988). A correlation coefficient for angular variables. In *Statistical Theory and Data Analysis II*, K. Matusita (ed.), 349–364. North Holland, New York.

Jammalamadaka, S.R. and SenGupta, A. (2001). *Topics in Circular Statistics.* World Scientific, Singapore.

Jordan, M.I. (2004). Graphical models. *Statist. Science* **19**, 140–155.

Juang, B.H. and Rabiner, L.R. (1991). Hidden Markov models for speech recognition. *Technometrics* **33**, 251–272.

Kelly, F.P. (1979). *Reversibility and Stochastic Networks.* Wiley, Chichester.

Kennedy, J.S. (1992). *The New Anthropomorphism.* Cambridge University Press, Cambridge.

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Rev. Econ. Studies* **65**, 361–393.

Koski, T. (2001). *Hidden Markov Models for Bioinformatics.* Kluwer Academic Publishers, Dordrecht.

Lange, K. (1995). A quasi-Newton acceleration of the EM algorithm. *Statistica Sinica* **5**, 1–18.

Lange, K. (2002). *Mathematical and Statistical Methods for Genetic Analysis*, second edition. Springer, New York.

Lange, K. (2004). *Optimization.* Springer, New York.

Lange, K. and Boehnke, M. (1983). Extensions to pedigree analysis V. Optimal calculation of Mendelian likelihoods. *Hum. Hered.* **33**, 291–301.

Le, N.D., Leroux, B.G. and Puterman, M.L. (1992). Reader reaction: Exact likelihood evaluation in a Markov mixture model for time series of seizure counts. *Biometrics* **48**, 317–323.

Leisch, F. (2004). FlexMix: A general framework for finite mixture models and latent class regression in R. *J. Statistical Software* **11**. http://www.jstatsoft.org/v11/i08/.

Leroux, B.G. and Puterman, M.L. (1992). Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* **48**, 545–558.

Levinson, S.E., Rabiner, L.R. and Sondhi, M.M. (1983). An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *Bell System Tech. J.* **62**, 1035–1074.

Lindsey, J.K. (2004). *Statistical Analysis of Stochastic Processes in Time.* Cambridge University Press, Cambridge.

Lindsey, J.K. (2008). **R** package 'Repeated'. URL http://popgen.unimaas.nl/˜jlindsey/rcode.html, accessed 27 July 2008.

Linhart, H. and Zucchini, W. (1986). *Model Selection.* Wiley, New York.

Little, R.J.A. and Rubin, D.B. (2002). *Statistical Analysis with Missing Data*, second edition. Wiley, New York.

Lloyd, E.H. (1980). *Handbook of Applicable Mathematics, Vol. 2: Probability.* Wiley, New York.

Lystig, T.C. and Hughes, J.P. (2002). Exact computation of the observed information matrix for hidden Markov models. *J. Comp. Graphical Statist.* **11**, 678–689.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, second edition. Chapman & Hall, London.

MacDonald, I.L. and Raubenheimer, D. (1995). Hidden Markov models and animal behaviour. *Biometrical J.* **37**, 701–712.

MacDonald, I.L. and Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete-valued Time Series.* Chapman & Hall, London.

McFarland, D. (1999). *Animal Behaviour: Psychobiology, Ethology and Evolution*, third edition. Longman Scientific and Technical, Harlow.

McLachlan, G.J. and Krishnan, T. (1997). *The EM Algorithm and Extensions.* Wiley, New York.

McLachlan, G.J. and Peel, D. (2000). *Finite Mixture Models.* Wiley, New York.

Mira, A. (2000). Exuviae eating: a nitrogen meal? *J. Insect Physiol.* **46**, 605–610.

Munoz, W.P., Haines, L.M. and van Gelderen, C.J. (1987). An analysis of the maternity data of Edendale Hospital in Natal for the period 1970–1985. Part 1: Trends and seasonality. Internal report, Edendale Hospital.

Newton, M.A., and Raftery, A.E., (1994). Approximate Bayesian inference with the weighted likelihood bootstrap (with discussion). *J. Roy. Statist. Soc.* B **56**, 3–48.

Nicolas, P., Bize, L., Muri, F., Hoebeke, M., Rodolphe, F., Ehrlich, S.D., Prum, B. and Bessières, P. (2002). Mining *Bacillus subtilis* chromosome heterogeneities using hidden Markov models. *Nucleic Acids Res.* **30**, 1418–1426.

Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007). Stochastic volatility with leverage: fast and efficient likelihood inference. *J. Econometrics* **140**, 425–449.

Pearl, J. (2000). *Causality: Models, Reasoning and Inference.* Cambridge University Press, Cambridge.

Pegram, G.G.S. (1980). An autoregressive model for multilag Markov chains. *J. Appl. Prob.* **17**, 350–362.

R Development Core Team (2008). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Raftery, A.E. (1985a). A model for high-order Markov chains. *J. Roy. Statist. Soc.* B **47**, 528–539.

Raftery, A.E. (1985b). A new model for discrete-valued time series: autocorrelations and extensions. *Rassegna di Metodi Statistici ed Applicazioni* **3–4**, 149–162.

Raftery, A.E., Newton, M.A., Satagopan, J.M. and Krivitsky, P.N. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity (with discussion). In *Bayesian Statistics 8*, J.M. Bernardo, M.J. Bayarri, J.O. Berger, A.P. Dawid, D. Heckerman, A.F.M. Smith and M. West (eds.), 1–45. Oxford University Press, Oxford.

Raftery, A.E. and Tavaré, S. (1994). Estimation and modelling repeated patterns in high order Markov chains with the mixture transition distribution model. *Appl. Statist.* **43**, 179–199.

Raubenheimer, D. and Barton Browne, L. (2000). Developmental changes in the patterns of feeding in fourth- and fifth-instar *Helicoverpa armigera* caterpillars. *Physiol. Entomology* **25**, 390–399.

Raubenheimer, D. and Bernays, E.A. (1993). Patterns of feeding in the polyphagous grasshopper *Taeniopoda eques*: a field study. *Anim. Behav.* **45**, 153–167.

Richardson, S. and Green, P.J. (1997). On Bayesian analysis of mixtures with an unknown number of components (with discussion). *J. Roy. Statist. Soc.* B **59**, 731–792.

Robert, C.P. and Casella, G. (1999). *Monte Carlo Statistical Methods.* Springer, New York.

Robert, C.P., Rydén, T. and Titterington, D.M. (2000). Bayesian inference in hidden Markov models through the reversible jump Markov chain Monte Carlo method. *J. Roy. Statist. Soc.* B **62**, 57–75.

Robert, C.P. and Titterington, D.M. (1998). Reparameterization strategies for hidden Markov models and Bayesian approaches to maximum likelihood estimation. *Statist. and Computing* **8**, 145–158.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* **23**, 470–472.

Rossi, A. and Gallo, G.M. (2006). Volatility estimation via hidden Markov models. *J. Empirical Finance* **13**, 203–230.

Rydén, T., Teräsvirta, T. and Åsbrink, S. (1998). Stylized facts of daily returns series and the hidden Markov model. *J. Appl. Econometr.* **13**, 217–244.

Schilling, W. (1947). A frequency distribution represented as the sum of two Poisson distributions. *J. Amer. Statist. Assoc.* **42**, 407–424.

Schimert, J. (1992). A high order hidden Markov model. Ph.D. dissertation, University of Washington.

Scholz, F.W. (2006). Maximum likelihood estimation. In *Encyclopedia of Statistical Sciences*, second edition, S. Kotz, N. Balakrishnan, C.B. Read, B. Vidakovic and N.L. Johnson (eds.), 4629–4639. Wiley, Hoboken, NJ.

Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice and Visualization*. Wiley, New York.

Scott, S.L. (2002). Bayesian methods for hidden Markov models: Recursive computing in the 21st century. *J. Amer. Statist. Assoc.* **97**, 337–351.

Scott, S.L., James, G.M. and Sugar, C.A. (2005). Hidden Markov models for longitudinal comparisons. *J. Amer. Statist. Assoc.* **100**, 359–369.

Shephard, N.G. (1996). Statistical aspects of ARCH and stochastic volatility. In *Time Series Models: In econometrics, finance and other fields*, D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (eds.), 1–67. Chapman & Hall, London.

Sibly, R.M. and McFarland, D. (1976). On the fitness of behaviour sequences. *American Naturalist* **110**, 601–617.

Silverman, B.W. (1985). Some aspects of the spline smoothing approach to nonparametric regression curve fitting (with Discussion). *J. Roy. Statist. Soc.* B **47**, 1–52.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.

Simpson, S.J. (1990). The pattern of feeding. In *A Biology of Grasshoppers*, R.F. Chapman and T. Joern (eds.), 73–103. Wiley, New York.

Simpson, S.J. and Raubenheimer, D. (1993). The central role of the haemolymph in the regulation of nutrient intake in insects. *Physiol. Entomology* **18**, 395–403.

Singh, G.B. (2003). Statistical Modeling of DNA Sequences and Patterns. In *Introduction to Bioinformatics: A Theoretical and Practical Approach*, S.A. Krawetz and D.D. Womble (eds.), 357–373. Humana Press, Totowa, NJ.

Smyth, P., Heckerman, D. and Jordan, M.I. (1997). Probabilistic independence networks for hidden Markov probability models. *Neural Computation* **9**, 227–269.

Speed, T.P. (2008). Terence's stuff: my favourite algorithm. *IMS Bulletin* **37(9)**, 14.

Spreij, P. (2001). On the Markov property of a finite hidden Markov chain. *Statist. Prob. Letters* **52**, 279–288.

Stadie, A. (2002). Überprüfung stochastischer Modelle mit Pseudo–Residuen. Ph.D. dissertation, Universität Göttingen.

Suster, M.L. (2000). Neural control of larval locomotion in *Drosophila melanogaster*. Ph.D. thesis, University of Cambridge.

Suster, M.L., Martin, J.R., Sung, C. and Robinow, S. (2003). Targeted expression of tetanus toxin reveals sets of neurons involved in larval locomotion in *Drosophila*. *J. Neurobiology* **55**, 233–246.

Timmermann, A. (2000). Moments of Markov switching models. *J. Econometrics* **96**, 75–111.

Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985). *Statistical Analysis of Finite Mixture Distributions*. Wiley, New York.

Toates, F. (1986). *Motivational Systems*. Cambridge University Press, Cambridge.

Turner, R. (2008). Direct maximization of the likelihood of a hidden Markov model. *Computat. Statist. & Data Analysis* **52**, 4147–4160.

van Belle, G. (2002). *Statistical Rules of Thumb*. Wiley, New York.

Visser, I., Raijmakers, M.E.J. and Molenaar, P.C.M. (2002). Fitting hidden Markov models to psychological data. *Scientific Programming* **10**, 185–199.

Viterbi, A.J. (1967). Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. *IEEE Trans. Inform. Th.* **13**, 260–269.

Wasserman, L. (2000). Bayesian model selection and model averaging. *J. Math. Psychology* **44**, 92–107.

Weisberg, S. (1985). *Applied Linear Regression*, second edition. Wiley, New York.

Welch, L.R. (2003). Hidden Markov models and the Baum–Welch algorithm. *IEEE Inform. Soc. Newsl.* **53**, pp. 1, 10–13.

Whitaker, L. (1914). On the Poisson law of small numbers. *Biometrika* **10**, 36–71.

Wittmann, B.K., Rurak, D.W. and Taylor, S. (1984). Real-time ultrasound observation of breathing and body movements in foetal lambs from 55 days gestation to term. Abstract presented at the XI Annual Conference, Society for the Study of Foetal Physiology, Oxford.

Yu, J. (2005). On leverage in a stochastic volatility model. *J. Econometrics* **127**, 165–178.

Zeger, S.L. and Qaqish, B. (1988). Markov regression models for time series: a quasi-likelihood approach. *Biometrics* **44**, 1019–1031.

Zucchini, W. (2000). An introduction to model selection. *J. Math. Psychology* **44**, 41–61.

Zucchini, W. and Guttorp, P. (1991). A hidden Markov model for space-time precipitation. *Water Resour. Res.* **27**, 1917–1923.

Zucchini, W. and MacDonald, I.L. (1998). Hidden Markov time series models: some computational issues. In *Computing Science and Statistics* **30**, S. Weisberg (ed.), 157–163. Interface Foundation of North America, Inc., Fairfax Station, VA.

Zucchini, W. and MacDonald, I.L. (1999). Illustrations of the use of pseudo-residuals in assessing the fit of a model. In *Statistical Modelling. Proceedings of the 14th International Workshop on Statistical Modelling, Graz, July 19–23, 1999*, H. Friedl, A. Berghold, G. Kauermann (eds.), 409–416.

Zucchini, W., Raubenheimer, D. and MacDonald, I.L. (2008). Modeling time series of animal behavior by means of a latent-state model with feedback. *Biometrics* **64**, 807–815.