**McLachlan G, Peel, D A** 2000: *Finite mixture models.* New York: Wiley xxii + 420 pp., £64.50 ISBN 0 471 00626 2.

The publication in 1988 of McLachlan and Basford's *Micture Models: Inference and Applications to Clustering* heralded much interest in applications of mixture models, in developments such as multinomial, hidden Markov chain and hidden Markov random field models, and in the computer estimation of such models. McLachlan and Peel comment that about 40% of the 800 or so references in this new monograph have been published since 1995. Their aim is to draw the subject together in book form. They are to be congratulated on the extent of their achievement.

Their main foci of attention are the EM algorithm for maximum likelihood estimation (not forgetting identifiability), Bayesian estimation via Markov chain Monte Carlo methods and the assessment of an appropriate number of model components, in that order. Mixtures with a non-finite number of components receive little attention; non-standard mixtures (ones with degenerate components) are mentioned occassionally.

Chapter 1 begins with a brief discussion of the value of mixtures of distributions as flexible models and a whirlwind tour of areas of application. On page 3, before infinite or finite mixtures have yet been defined, attention is drawn to the influence of the EM algorithm. This is followed by an overview fo the contents of the book. Only then are we given a basic definition of a mixture model and comments are made about the interpretation of mixture models. Chapter 1 also includes sections, *inter alia*, about the shapes of some univariate normal mixtures, the modelling of asymmetric data, spurious clusters, sampling designs for classified data, parametric and nonparametric formulations, identifiability, clustering of data via mixture models, hidden Markov models, testing for the number of components and a brief history of mixture models. It ends with a section on notation (a similar section on acronyms would have been helpful).

I did not understand the ordering of the sections in Chapter 1. Throughout the rest of the book I found the presentation of the material much more coherent.

Chapter 2 takes us back to the EM algorithm as a method for the maximum likelihood fitting of mixture models. The choice of starting values and stopping criteria, the calculation of the observed information matrix and the estimation of standard errors by information-based methods and by boot-strapping are all covered.

In Chapter 3, the results in Chapter 2 are applied to mixtures of normal components; illustrative examples are provided. The authors remark,

> 'Given the tractability of the multivariate normal distribution, it is not surprising that mixture modeling of continuous data is invariably [*sic*] undertaken by normal mixtures'.

The Bayesian approach to the fitting of mixture models by using posterior simulation via Markov chain Monte Carlo methods is the topic in Chapter 4. Its advantages and disadvantages are discussed and methods such as the use of partially proper priors are put forward for dealing with problems like improper posterior distributions. Label switching is also addressed.

Chapter 5 is about the fitting of mixture models with non-normal component densities (continuous, categorical and mixed continuous–categorical). The maximum likelihood fitting of discrete components (Poisson, and also binomial) is examined within the wider context of a mixture of generalized linear models. Mixtures of generalized linear models for handling overdispersion, the mixtures-of-experts model an dthe hierarchical mixtures-of-experts model are considered.

By this state of reading the book I felt that it should have a subtitle indicating its emphasis on methods of model fitting. In Chapter 6, attention moves away to the choice of the appropriate number of components (order) of a mixture model. Two approaches, penalized log-likelihood and hypothesis testing using the likelihood ratio statistic, are covered.

The author's recent work on fitting mixtures of (multivariate) $t$-distributions is the topic in Chapter5 7. This yields a longer-tailed, and hopefully more robust, alternative to fitting mixtures of normals. Chapters 8–10 are on mixtures of factor analysers, fitting finite mixture models to binned and truncated multivariate data (via the EM algorithm) and the use of mixtures to model failure time data. In Chapter 11, there is a case-study illustrating the use of mixture models in the analysis of multivariate directional data. Chapter 12 studies various methods for improving the speed of the EM algorithm to make it useful for large databases. Finally, Chapter 13 introduces some recent advances concerning hidden Markov chain models (for the one-dimensional case) and hidden Markov random fields (for two or higher dimensions).

Throughout the book nearly 30 (real) data sets and some synthetic ones are used to illustrate the methodology. Those that are in the public domain are obtainable from

http://www.maths.uq.edu.au/ gjm

Appendix gives a brief description of the authors' EMMIX program and other software for fitting mixture models; it includes details of their Web sites. The *circa* 800 papers and books listed at the end of the book seem to be mentioned, albiet sometimes very briefly, somewhere in the book. Their comprehensiveness means that they form a very good, recent bibliography.

This book resembles an encyclopaedia. From Chapter 2 onwards, the chapters are like articles—they are largely stand alone and nearly all give an encyclopaedic coverage of a particular aspect of the fitting of mixture models. Encyclopaedias often give no indication of their putative readership. Not do these authors. If you are the sort of person who needs to own this book, then you probably know about it already. If you are not, then you should realize that this is an important area of statistics that currently commands much interest. Try to persuade your library to buy a copy.

Adrienne W. Kemp
St Andrews University