Chapter 5 describes, at some length, two-stage estimation: this is feasible when there is enough data per individual to form reliable estimates of individual regression parameters; these can then be used as input data, accompanied by appropriate variability estimates, to fit the between-individuals model.

Chapter 6 addresses the case where the within-individuals data sets are not large. The main theme is linear approximation of the nonlinear functions in order to apply the numerical techniques associated with linear models. Many complicated and ingenious wheezes are described to get round the problem of nonlinearity and the integrations involved. However, it should be borne in mind that these approximations are in addition to the asymptotics relied upon for inference.

Chapter 7 describes certain non- and semi-parametric methods. Section 7.2 goes into some theoretical detail concerning non-parametric maximum likelihood. A method referred to as 'smooth non-parametric' is presented, taken from an author's own joint research. This chapter, in particular, presents material that has not appeared widely before. As pointed out by the authors, weaker distributional specifications rely on larger sample sizes.

Chapter 8 gives a nice overview of Bayesian MCMC, focussing mainly on Gibbs sampling since full conditionals of parameter-sets are readily available in the present context.

Chapter 9 presents three case studies in detail from pharmacology, after giving a general introduction to the models involved. The discussion of modelling, fitting and results for these data sets is careful and informative. However, the analyses performed in this chapter, which follows chapter 8, are set in the longer-established mould of previous chapters.

Chapter 10 examines some of the issues which arise in analyzing drug assay data. The treatment is detailed and looks comprehensive enough to form a useful introduction to the field for potential practitioners.

Chapter 11, Further Applications, describes three more case studies, one each in crop science, forestry and seismology. The account brings out further practical aspects of modelling, fitting and non-Bayesian inference.

Chapter 12, Open Problems and Discussion, is an agenda for further developments.

Overall, the book gives a very well-written account of the field over the past few decades, focussing mainly on US work and including much of the authors' own, plus a glimpse of the future. It fairly reflects that literature over the years in dwelling at length on certain computational methods for maximum likelihood estimation. There is a leaning towards biopharmaceutical applications, this being a field in which the authors are acknowledged authorities. There are no exercises, but enough detail of the methodology is given, together with helpful guidance on available software, to enable the keen novice to try his hand. Lastly, I hope I'll be forgiven for just adding a couple of citations to those given in the book: a slightly different perspective on nonlinear regression models for repeated measures is described briefly in Crowder and Hand (1990, Section 9.4) and at greater length in Hand and Crowder (1995, chapter 8).

REFERENCES

Fletcher, R. and Powell, M. J. D. (1963). A rapidly convergent descent method for minimization. *Computer Journal* **6**, 163–168.

Crowder, M. J. and Hand, D. J. (1990). *Analysis of Repeated Measures.* London: Chapman & Hall.

Hand, D. J. and Crowder, M. J. (1995). *Practical Longitudinal Data Analysis.* London: Chapman & Hall.

Whittle, P. (1961). Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute* **39**, 1–26.

M. J. CROWDER
University of Surrey
Guildford, England.

CARROLL, R. J., RUPPERT, D., and STEFANSKI, L. A. **Measurement Error in Nonlinear Models.** Chapman & Hall, London, 1995. xxiv + 305 pp. £29.95/$54.95. ISBN 0-412-04721-7.

It has been a pleasure to review this book, which provides both an accessible, carefully motivated account of a range of models, illustrated by real examples, and detailed discussion of the statistical theory.

Seven examples are given in the first chapter, including nutrition studies, bioassay in a herbicide study, lung function in children and the atom bomb survivors data. General models and issues in measurement error analysis are then reviewed: functional and structural models, and replicate measurements, validation studies and instrumental variables as methods for assessing error in covariates. A brief tour of the book is provided, which ensures that the reader can select material appropriate to their interests. The 'lovely and appealing folklore' that measurement error always results in attenuation is shown to be false in Chapter 2, which summarises methods for linear models. Readers are referred to Fuller (1987) for details of linear measurement error models, but are left with no illusions about the complexity which can arise in measurement error problems.

The next six chapters are structured so that readers with a particular applied problem to consider can read an overview and basic description of a method before being presented with a worked example. Those wishing greater understanding of the theory can read the remainder of the chapter.

Functional models, in which no or minimal assumptions are made about the distribution of variables measured with error, are described in Chapters 3 to 6. Regression calibration, in which the error is adjusted by regression on a related variable, and other, accurate, covariates, is described in Chapter 3. The idea that the effect of measurement error on an estimator can be determined experimentally motivates simulation extrapolation in Chapter 4. Instrumental variables, which are correlated with the variable of interest, say X, independent of the error in X and independent of the outcome variable given the other variables, are described in the context of generalized linear models. Chapter 6 rounds off functional models by considering conditional-score and corrected-score methods which are fully consistent for important models such as linear, logistic and log-linear models.

Structural models, in which the distribution of the true predictor is modelled parametrically, are discussed in Chapters 7 and 8. Likelihood and quasi-likelihood are discussed first, based on the type and extent of data available. The link with an interpretation as a missing data problem is mentioned. Bayesian methods are briefly introduced in Chapter 8.

More specialised topics are addressed in the remaining six chapters. Semi-parametric methods when there are internal validation data, and unknown link functions in the context of functional techniques are described in Chapters 9 and (a seven-page) 10. Another short chapter, 11, summarises the effects of measurement error on hypothesis testing. Non-parametric estimation of densities and regression are reviewed in Chapter 11.

In most applications one thinks of the explanatory variable as measured with error, but the response variable might also be inaccurate, and this is considered in Chapter 13. A variety of topics, such as logistic case-control studies, survival analysis, and mixture methods as functional modelling are treated in the final chapter. The techniques underlying the theory are reviewed in an appendix. Programmes to implement various methods are available in Splus and SAS by email.

The authors acknowledged in the preface that the book necessarily reflects their interests. This modesty is hardly necessary, as this book gives a good account of the general ideas and strategies, and should stimulate others to expand research in an interesting and very important area of statistics. Diagnostics and model checking will repay further study. The range of the authors' appreciation of the actual issues and challenges pertaining to particular data sets, and extensive theoretical justification of estimation techniques is very impressive.

References

Fuller, W. W. (1987). *Measurement Error Models*. New York: Wiley.

J. L. Hutton
Department of Statistics
University of Newcastle upon Tyne
Newcastle upon Tyne, England.

HAND, D. J. **Construction and Assessment of Classification Rules**. Wiley, Chichester, England, 1997. xii + 214 pp. £34.95/$60.00. ISBN 0-471-96583-9.

This is a complete rewrite of the author's earlier book on the problem of 'supervised classification', where we have a multivariate data set with known group memberships and wish to construct rules for deciding the group membership of new observations. This is an old statistical problem but, due to the influence of computers and computer scientists, is probably a more popular area of research and application than ever before. An introductory level book summarising the state of play in the field is therefore very welcome, especially one written by someone with many publications in the area.

There are at least four themes running throughout the book:

(i) In expert hands and in an appropriate context all of the methods described can perform well. The choice of method is highly context-specific and appropriate advice can only concern characteristics of the methods – global recommendations do not make sense. Choice of the way(s) of assessing a derived rule are similarly context-specific.

(ii) The classic 'bias versus variance' conflict. Having a very flexible method for constructing rules may seem like a good idea, but often leads to overfitting and modelling of merely the random variation in the design set.

(iii) The relationship between classification and regression. The former can be viewed as an (important) special case of the latter, leading to many fruitful interconnections.

(iv) The similarities and differences between the work done by the statistical and the computer science (machine learning, expert systems) communities. Most notably, computer scientists usually deal with categorical variables and often assume perfect separability between groups, that is that two observations with the same values on all variables must belong to the same group. Statisticians, growing up with the normal distribution, tend to prefer continuous distributions and would probably never even consider assuming perfect separability. Statisticians also place far greater emphasis on rigorous estimation procedures whereas 'machine learning' tends to be rather more *ad hoc*, often considering one observation at a time. One can't help feeling that the author believes that computer scientists have much more to learn from statisticians than vice versa, and agreeing with him, but perhaps a computer scientist reading the book would not agree. In particular, he concludes that the statisticians' approach is much more appropriate in 'natural' (e.g. biological) applications.

Chapters 1–2 are a good exposition of the whys and wherefores, covering the basic ideas plus the classical Fisher approach and its extensions.