provide only a superficial discussion of the random and mixed effects analy-
ysis of variance” (p. ix). It is intended as a text for an introductory analysis of
variance service course and requires little more than a precalculus understand-
ing of probability, estimation, and hypothesis testing. Analysis of variance as
covered in this text is the decomposition of a total sum of squares into addi-
tive components in classification models. The statistical models of concern are
those for the one-way, two-way crossed, three-way crossed, two-way nested,
three-way nested, and partially nested classifications. Analysis of covariance or
variability partitioning in regression models are not part of the contents.
The statistician will find little new material beyond well-established texts such as
Cochran and Cox (1957), Scheffé (1959), Searle (1971, 1987), and Searle,
Casella, and McCulloch (1992), to name just a few of the important works in
this area. The student and practitioner will benefit from a well-balanced mix-
ture of statistical theory, formulas, and explanations and the great care exer-
cised by the authors in discussing properties and analysis of fixed, random,
and mixed models in parallel.

The main chapters of the book follow the previously mentioned list of
models and are organized into similar topics throughout. A chapter is devoted to
a particular classification model, and two-way models with and without
interactions are discussed separately. A discussion of model assumptions is
followed by the sum of squares partition, mean squares and their expectations,
sampling distribution of mean squares, ANOVA F test construction and test
power, point estimation of means and variance components, interval estima-
tion, multiple comparison methods, and departures from model assumptions.
The one-way classification is discussed in extra detail, covering 113 pages
after a brief introductory chapter. Distributional properties of statistics are
examined under normality throughout, not randomization theory.

Where differences exist between fixed, random, and mixed models, these
are discussed in separate subsections of the same topic. The reader is thus
allowed a direct comparison of the three model types in terms of, for example,
equivalent mean squares and their sampling distributions. In contrast to many
introductory textbooks covering analysis of variance, the reader is not left with
the impression that the most common is with fixed effects ANOVA, whereas
random and mixed models are of lesser importance. Each chapter contains
worked examples for a fixed, random, and (where applicable) mixed model.
The data are analyzed with and without statistical packages (SAS, SPSS, and
BMDP) using small datasets from standard textbooks, such as Sokal and Rohlf
(1995) and Steel and Torrie (1980), or artificial data. Similar data sources are
used for exercises that mark the conclusion of each chapter. The number of
exercises per chapter varies from 1 ( Chap. 9, “Finite Population and Other
Models”) to 27 ( Chap. 10, “Some Simple Experimental Designs”). Answers
to exercises are not provided.

The book employs several devices to aid readability for the target audience.
Instead of linear algebra tools, the discussion relies on summation formulas to
demonstrate sum of squares partitioning and expected mean squares deriv-
ation. The body of the text contains an abundance of remarks and footnotes
distinguished from the main text by a smaller type. These additions point the
reader to additional literature and provide extra detail that the authors feel
can be glossed over by less interested readers. The remarks are frequently
more than half a page long and contain very pertinent information, insights,
and important details this reviewer feels all readers should be
exposed to. The remarks also contain a large number of the more than 630
references Sahai and Ageel assembled. Discussing additional references in
a separate section at the end of a chapter and merging the remarks into the main
body of the text would have strengthened the text in my opinion.

Computer implementation is dealt with in two ways. For each of the
worked examples in the main chapters, SAS, SPSS, and BMDP program-
ning statements and output are shown to produce the analysis of vari-
ance and F tests. Statements for mean separation, contrasts, or simple
effect tests are not provided. The final chapter discusses analysis of vari-
ance using these three packages in about 25 pages. For the worked exam-
oples of the main chapters, computer input and output are shown in boxes
side by side, with the input appearing on the left. The output is not anno-
tated and little help is provided to reconcile the computer results with
earlier hand calculations. Due to the side-by-side format of the computer
input-output, some of the SAS output tends to swap over into the input
box, making it difficult for the novice user to determine where program-
mation statements end and results begin. For higher-order models, the box
allotted for the programming statements is too small. Individual statements
then occupy several lines (up to five) and are difficult to read. Concerning
SAS, procedures glm and aov receive particular attention; nested models
are analyzed with proc glm and proc nested. Surprisingly, proc mixed
receives very little attention; not a single worked example employs this pro-
cedure.

A comprehensive text that covers analysis of variance for fixed, random,
and mixed models in tandem at a mathematical level appreciable by non-
statistics majors is a wonderful tool. However, many service courses will not
be dedicated in their entirety to just analysis of variance and will have either
a sizeable regression or experimental design component or both. Analysis of
variance methods in linear regression analysis are not discussed in this text.
Experimental designs are discussed in a separate chapter toward the end of
the text ( Chap. 10, pp.). Only the classical designs are discussed in some
detail. Worked examples are provided with simple treatment structures for
these designs up to the standard split-plot design. Experimental designs
where mixed model ideas are particularly pertinent, such as incomplete block,
crossover designs, and repeated measures designs, receive only brief attention.
A book concerned with mixed classification models should discuss the dif-
fferences between inter- and intrablock analyses in unbalanced or incomplete
block designs.

In summary, this text is a valuable source on analysis of variance in clas-
sification models for the practitioner and nonstatistics major that covers fixed,
random, and mixed models equally. It is well organized and well written at a
difficulty level that precisely meets the target audience’s needs.

O. Schabenberger
Virginia Tech

REFERENCES
Cochran, W. G., and Cox, G. M. (1957), Experimental Designs, New York:
Wiley.
Scheffé, H. (1959), The Analysis of Variance, New York: Wiley.
Searle, S. R. (1971), Linear Models, New York: Wiley.
Searle, S. R., Casella, G., and McCulloch, C. E. (1992), Variance Compo-
nents, New York: Wiley.
Sokal, R. R., and Rohlf, F. J. (1995), Biometry, New York: Freeman.
Steel, G. D., and Torrie, J. H. (1980), Principles and Procedures of Statistics:
A Biometrical Approach, New York: McGraw-Hill.

Epidemiology: Study Design and Data Analysis.

Mark Woodward. Boca Raton, FL: CRC Press, 1999. ISBN 1-5848-8009-0. xiv +
699 pp. $79.95.

The orientation and intended users of this text are explained in the first
two sentences of the preface: “This book is about the quantitative aspects
of epidemiological research. I have written it with two audiences in mind:
the researcher who wishes to understand how statistical principles and tech-
niques may be used to solve epidemiological problems, and the statistician
who wishes to find out how to apply his or her subject in this field.”

The titles and content summaries of the 11 chapters give a good overview
of the breadth of topics covered:

1. Fundamental Issues: case studies; populations and samples; measuring
disease; measuring risk factors; causality; studies using routine data; study
design; data analysis.
2. Basic Analytical Procedures: types of variables; tables and charts;
inferential techniques for categorical variables; descriptive techniques
for quantitative variables; inferences about means; inferential techniques for
non-normal data; measuring agreement; assessing diagnostic tests.
3. Assessing Risk Factors: risk and relative risk; odds and odds ratio;
prevalence studies; testing association; attributable risk; rate and relative rate.
4. Confounding and Interaction: The concept of confounding; identi-
fication of confounders; assessing confounding; standardization; Mantel–
Haenszel methods; the concept of interaction; testing for interaction; dealing
with interaction.
5 Cohort Studies: design considerations; analytical considerations; cohort
life tables; Kaplan–Meier estimation; comparison of survival probabilities;
person-years; period-cohort analysis.
6. Case-Control Studies: basic design concepts; basic methods of analy-
sis; selection of cases; selection of controls; matching; the analysis of matched
studies.
7. Intervention Studies: ethical considerations; avoidance of bias; parallel group studies; crossover studies; sequential studies; allocation to treatment group.

8. Sample Size Determination: power; testing a mean value; testing a difference between means; testing a proportion; testing a relative risk; case-control studies.

9. Modelling Quantitative Outcome Variables: statistical models; one categorical explanatory variable; one quantitative explanatory variable; two categorical explanatory variables; model building; general linear models; several categorical variables; model checking; confounding; nonnormal alternatives.

10. Modelling Binary Outcome Data: problems with standard regression models; logistic regression; interpretation of logistic regression coefficients; generic data; multiple logistic regression models; tests of hypotheses; confounding; interaction; model checking; case-control studies; outcomes with several ordered levels.

11. Modelling Followup Data: basic functions of survival time; estimating the hazard function; probability models; proportional hazards regression models; the Cox proportional hazards model; the Weibull proportional hazards model; model checking; Poisson regression.

The numerous explanations, examples, and references for analytical techniques are a major strength of the book. For instance, the example for general linear models (Example 9.11) is several pages long and includes a detailed discussion of analysis of variance tables, Type I and III sums of squares, and parameter estimates obtained from SAS PROC GLM. The SAS code used to run this example is presented in Appendix A, along with the SAS code for 40 other examples in the book. Such thorough details are provided for most methods covered in the book, and datasets for all examples are either given in Appendix C or can be downloaded from a web site. Specific analyses can be practiced by working the exercises at the end of each chapter, all of which have solutions provided.

SAS is the only software package for which procedures and output are presented and discussed, although other packages available for a given technique are generally mentioned. Such heavy reliance on SAS (specifically, SAS 6.11) may result in this aspect of the book becoming outdated in the future, but it is in keeping with the practical, applied approach used by Woodward. One aspect that already seems outdated is the use of PROC GENMOD for logistic regression. This may be a practice left from before PROC LOGISTIC became available, because the latter more readily provides odds ratios and confidence intervals (at least for terms not involved in interactions) as well as model diagnostics and goodness-of-fit tests.

The writing style is quite clear and straightforward, although there are occasions where it is a bit too elementary. One would expect readers of a book such as this to have no trouble with the logic of the sentence. “Binary variables are, as with other categorical variables, generally summarized by proportions or the equivalent percentages, that is, proportions multiplied by 100” (p. 47). In fact, much of Chapter 2 could have been omitted. It is apparently included so that the book will be “self-contained,” even though the author has “assumed that the reader has some basic knowledge of statistics” (Preface, p. xiii). The quest to cover all basic statistics in a single chapter leads to some oversimplifications, such as “The p value is the exact probability of getting a result as extreme as that observed for the test statistic when the null hypothesis is true” (p. 47).

It is ambitious to attempt fully covering both epidemiologic study design and analytical methods, even in 700 pages. This goal is not quite achieved, because the depth of discussion is less than ideal for a few issues. For example, there is no mention of intention-to-treat analysis, but it is implied that the opposite approach is preferred, stating “Compliance is recorded because we would normally exclude from the analysis any patient who has deviated importantly from the rules of the protocol” (p. 297). Such exclusions nullify the benefits of randomization by potentially reintroducing confounding differences between the groups (see, e.g., Friedman, Furberg, and DeMets 1985; Hulley and Cummings 1988). There is also no mention of some practical issues such as ways to enhance compliance among subjects in intervention studies.

The chapter on cohort studies does not cover strategies for following a cohort through time. For case-control studies, selection of cases and controls is presented, but measurement of exposures and confounders is not. The distinction between differential and nondifferential misclassification is not made, and the biasing effects of nondifferential misclassification are not explained. There is discussion of the need for the rare disease assumption for odds ratios to approximate relative risks, but no discussion of the lack of need for this assumption under some case-control sampling designs (Rodrigues and Kirkwood 1990) or discussion of these sampling designs (Rothman and Greenland 1998).

The guidelines for identifying confounders ( Sect. 4.3) are somewhat more deterministic than is often feasible, being based on “a priori knowledge of the supposed biological processes at work” (p. 152). If we have such knowledge, we would often not need to conduct the study. If we lack such knowledge, how do we decide whether to control for a factor or not? [One example in this section is difficult to follow because of omission of the word “Not” at the start of the sentence “Eating fruit causes a low intake of vitamin C which causes scurvy” (p. 151).]

Most analytical topics that are mentioned but not described in detail are thoroughly referenced. When discussing intervention studies, for example, Woodward presents sequential studies, points out that analysis of such studies is complex and beyond the scope of the book, and provides references for further information on the topic. Likewise, in the introduction to the chapter on sample size, many references are given for where to find discussion of sample size issues beyond testing of means, proportions, relative risks, and odds ratios.

There are a few analytical details that are not clearly described. When discussing comparison of standardized mortality ratios (or “standardized event ratios”) in Section 5.6.3, no mention is made that such comparisons are not strictly valid unless the age structures of the groups being compared are identical (Breslow and Day 1987). In the section on 2×2 crossover studies, the author states that if there is a significant treatment by period interaction, “we would be forced to use only the data from the first period of treatment and analyse as a parallel group study” (p. 308). This approach has been shown to produce a biased test statistic, with an actual significance level for treatment effect often much higher than that computed (Freeman 1989).

For logistic regression, although there is much discussion of examining the relative fit of nested models, there is little discussion of examining the absolute fit of the logistic model, with no mention of, for instance, the Hosmer-Lemeshow goodness-of-fit test [and also, surprisingly, no mention of Hosmer and Lemeshow’s (1989) Applied Logistic Regression, which is probably the standard reference on the subject, at least on this side of the Atlantic]. When introducing conditional logistic regression for matched data, the author notes that in “the special case of 1:1 matching the conditional likelihood can be reduced to the standard likelihood, and thus standard logistic regression procedures can be used (see Collett, 1991)” (p. 499). This does not make it clear that the data must first be modified, because employing an unconditional logistic regression procedure on raw data from a 1:1 matched study will not produce the correct conditional maximum likelihood estimates of coefficients.

These are mostly minor concerns. The book has a number of features, both major and minor, that I greatly appreciate. Major ones, in addition to those discussed earlier, include the detailed sections on checking model assumptions for each model presented and the chapter on modeling quantitative outcomes, a topic that is often ignored in discussions of epidemiologic analyses. Minor ones include the suggestion to refrain from terminology that is sometimes employed in epidemiology: use of the term “relative risk” when an odds ratio has been computed and use of the term “multivariate” for multivariable models with one dependent variable. Also well explained is the lack of standard nomenclature for terms such as “attributable risk.”

This book covers an impressive amount of material, including most of the important topics regarding epidemiologic study design and analysis. The clear explanations and thorough examples of analytical techniques make it well suited as a textbook and reference work for analysis of epidemiologic studies.

Peter Callas
University of Vermont

REFERENCES

Breslow, N. E., and Day, N. E. (1987), Statistical Methods in Cancer Research: Volume II. The Design and Analysis of Cohort Studies, Lyon: International Agency for Research on Cancer.

Freeman, P. R. (1989), “The Performance of the Two-Stage Analysis of Two-Treatment, Two-Period Crossover Trials,” Statistics in Medicine, 8, 1421–1432.

Friedman, L. M., Furberg, C. D., and DeMets, D. L. (1985), Fundamentals of Clinical Trials (2nd ed.), Littleton, MA: PSG Publishing Company.
Hosmer, D. W., and Lemeshow, S. (1989), Applied Logistic Regression, New York: Wiley.
Hulley, S. B., and Cummings, S. R. (1988), Designing Clinical Research, Baltimore: Williams and Wilkins.
Rodrigues, L., and Kirkwood, B. R. (1990), "Case-Control Designs in the Study of Common Diseases: Updates on the Demise of the Rare Disease Assumption and the Choice of Sampling Scheme for Controls," International Journal of Epidemiology, 19, 205–213.
Rothman, K. J., and Greenland, S. (1998), Modern Epidemiology (2nd ed.), Philadelphia: Lippincott-Raven.

Analysis of Health Surveys.

Edward L. Korn and Barry I. Graubard. New York: Wiley, 1999. ISBN 0-471-13773-1. xii + 382 pp. $84.95. The analytic components of books about survey sampling often restrict their discussion to the analysis of relatively simple parameters of interest, such as population means and linear regression slopes. On the other hand, texts not devoted to survey sampling rarely discuss in detail the effect of complex sample designs on data analysis. This book provides a great service to the health research community by tying together the tools of modern statistical analysis and survey research in one package. Korn and Graubard write clearly and concisely, using well-chosen examples to illuminate potentially confusing concepts.

The book has nine chapters: 1, Introduction; 2, Basic Survey Methodology; 3, Statistical Analysis with Survey Data; 4, Weights and Impu- tation; 5, Additional Issues in Variance Estimation; 6, Cross-Sectional Analyses; 7, Analysis of Longitudinal Surveys; 8, Analysis Using Multiple Surveys; and 9, Population-Based Case-Control Studies. The first five chapters are methodological, whereas the last four apply the methodology to specific examples of familiar large health-related surveys, such as the National Health Interview Survey and the National Health and Nutrition Examination Survey.

Chapter 2 briefly summarizes single-stage sampling plans, including systematic, stratified, and probability-proportional-to-size sampling, and their combination into multistage sampling typical of health surveys. Alternatives such as capture-recapture designs are also discussed briefly. The authors rightly defer detailed discussion of sample design, but further background from classic texts such as Kish (1966), Cochran (1977), or Thompson (1992) may be of use to the student or interested reader. Whereas the variance estimates obtained under the assumption of simple random sampling will often be unbiased for complex sample designs, this chapter also discusses the use of Taylor series approximations and replication techniques such as balanced repeated replication, jackknifing, and bootstrapping. As in the remainder of the book, justification for these methods is provided primarily through sketches of proofs and examples; those interested in more formal presentation are referred elsewhere (e.g., Wolter 1985). Using auxiliary population information to adjust sample weights (poststratification) and sources of error, such as nonresponse and sampling frame undercoverage, are also discussed.

Chapter 3 discusses estimation of and inference about descriptive statistics such as population means and totals, as well as the use of linear regression, binomial and multinomial logistic regression, and proportional hazard regression models. The authors also emphasize nonparametric estimation, in particular, kernel density estimation to summarize univariate distributions and cubic splines to summarize bivariate associations or extend regression models. The authors assume previous background in these areas and focus on their application in a complex sample design setting. Straightforward extensions of diagnostics such as added variable plots are also described in detail. A hypothesis test of $H_0: \beta = \beta_0$ versus $H_1: \beta \neq \beta_0$, for a vector of $q$ regression parameters using the Wald test statistic $W = (\hat{\beta} - \beta_0)' V^{-1} (\hat{\beta} - \beta_0)$ compared to a $\chi^2_q$ distribution typically has poor control of type I error in most survey settings because of the instability in $V$, the estimated covariance matrix of $\hat{\beta}$ (Thomas and Rao 1987). Korn and Graubard suggest using the general purpose correction $d - q + 1) W/dq$ (Rao and Thomas 1988), under which the null hypothesis may be compared to an $F_{q, d-q+1}$ distribution, where $d$ is the degrees of freedom available for variance estimation, typically the number of primary sampling units (PSU’s) minus the number of strata in a multistage sample design.

Because sampled units in surveys have unequal probabilities of selection, weights equal to the inverse of the probability of selection are often used to reduce or remove bias in estimators of interest. Chapter 4 describes construction and use of survey weights. Because use of weights often entails an increase in the variance of the estimator, a choice must often be made about whether or not the reduction in bias is worth the trade-off in increased variance. The authors neatly show this trade-off in the case of simple linear regression of a $Y$ on $X$. The unweighted least squares estimate of the regression slope is unbiased if the relationship between $Y$ and $X$ is truly linear; if the relationship is nonlinear, the weighted least squares estimate corrects for the model misspecification to yield an unbiased estimate of the least square regression slope in the population. Hence the need to use sample weights in this setting is not always clear; if the model is correctly specified, their use is superfluous and only adds variance; if the model is badly misspecified, the linear slope may be a meaningless target quantity. The authors provide an inefficiency measure under the null hypothesis that the unweighted estimator is unbiased and use this as a guidepost for deciding whether or not to utilize sample weights, depending on the type of analysis being conducted.

The use of weights in this text is discussed from the standard design-based approach to estimation and inference, in reference to a population parameter of interest. Here additional discussion of the model-based approach to analysis of probability-of-inclusion-weighted survey data might have been useful. For example, use of random-effects models to borrow strength across post-strata when the poststratified cell counts are small (Little 1993; Lazzeroni and Little 1998) provides an alternative to the either-or choice of fully weighted or unweighted analysis.

Chapter 4 also discusses imputation of missing data. The authors describe mean and hot-deck imputation of item-level missing values as a means to improve inference through both reduced bias and variance, and describe the multiple hot-deck imputation of Rubin and Schenker (1986) as a method for correctly adjusting the variance of an estimator to account for the uncertainty in imputation. A novice reader of this text might not fully grasp the need for such a correction; additional discussion of the Bayesian approach of multiple imputation (Rubin 1987; Shafer 1997) or frequentist use of correction factors in single imputation (Rao and Shao 1992; Rao 1996) would have been helpful.

Chapter 5 continues the discussion of variance estimation in Chapter 2. Because the Wald test will be extremely inefficient (or undefined) when $q$ is moderate or large, because the degrees of freedom are limited by the number of PSU’s minus the number of strata, alternative approaches are described, including testing components of $\beta$ one at a time using Bonferroni adjustments, correction of the simple random sampling estimate of the variance via Satterthwaite approximations, and additional replication techniques. Alternatively, different stages of the sample design may be ignored to increase $d$: Korn and Graubard describe a simple $F$ test for deciding when such approximations are reasonable. Unfortunately, the authors elected not to discuss the software packages such as commercial (SUDAAN, Stata) and free (VPLX, IVEWare) now available for variance estimation in complex sample design (see http://www.fas.harvard.edu/stats/survey-software/survey-soft.html). Although any discussion of software inevitably becomes dated rapidly, an instructor or interested reader will need to look elsewhere to learn what is available. A discussion of population versus superpopulation inference is also provided in this chapter.

Chapters 6–9 describe analyses of cross-sectional, longitudinal, and population-based case-control studies. Korn and Graubard discuss the choice of variables to be analyzed, describe their decision of whether or not to use sampling weights, and interpret their analyses and associated diagnostics in a creative, lively, and clear fashion. Teachers and practitioners of applied statistics alike will find these chapters useful, whether or not they regularly analyze survey data. For example, their use of predictive marginals to show the effect of interactions is an attractive alternative to traditional descriptions of interactions in any setting.

The authors include 30–40 superb problems following each methodological chapter, including questions about sample design, use of cubic splines, interpretation of model parameters in complex sample design settings, the bias of unweighted linear regression parameters, and so forth. The reader will need to work at least a subset of these problems to obtain the full value of the text.

This book would be an excellent textbook for those training epidemiologists or other nonstatisticians in the analysis of survey data. It could also serve as a main or secondary text in a master’s level biostatistics or an advanced undergraduate applied statistics course, and would be a convenient reference for any statistician who routinely deals with survey data.

Michael Elliott
University of Pennsylvania