CBCS
http://agrobiol.sggw.waw.pl/cbcs

Communications in Biometry and Crop Science
Vol. 1, No. 2, 2006, pp. 139–141

International Journal of the Faculty of Agriculture and Biology,
Warsaw Agricultural University, Poland

BOOK REVIEW

EXPERIMENTAL DESIGN AND DATA ANALYSIS FOR BIOLOGISTS by Gerry P. Quinn, Michael J. Keough. Cambridge University Press, Cambridge, UK. Published in 2002, reprinted with corrections in 2003, (paperback), price £31, 556 pp., ISBN 0 521 00976 6.

There are numerous books dealing with statistical methodology, both theoretical and applied, and a choice between them is not easy. For crop science, a statistical book should be concerned with situations that crop scientists commonly encounter during their research as well as with important issues that may be of particular interest to them. Quinn and Keough's book is an excellent guide for non-statisticians to simple and advanced issues of designing and analyzing biological experiments. Not only does the book help in designing experiments and analyzing the results, it also helps understand many intricate statistical problems that usually seem too advanced for non-statisticians, such as split-plot designs and repeated measures, complex analysis of variance, or multivariate methods.

The book provides the reader with answers to many questions related to design and analysis of biological experiments. Although most examples are related to ecological and biological studies, the book addresses most statistical issues that may be encountered in crop science. These are, among others, estimation and hypothesis testing, correlation and regression analysis (with reference to works on other methods of studying associations), analysis of variance—ANOVA (the classical analysis as well as analyses under more complex designs, e.g., split-plot, nested designs, and repeated measures), analysis of covariance, and multivariate analyses (e.g., multivariate ANOVA, discriminant analysis, classification methods, principal component analysis, and correspondence analysis). The problems are presented in a simple manner, without complex mathematical language (even though the authors are not afraid to employ it), but with many helpful examples. The reader will find the book interesting, comprehensive, and, last but not least, sufficient as a guide to simple and advanced statistical methods that may be applied in crop science.

Both authors are experts in the application of statistical methods in biology and ecology. Because of its high quality, the book has been cited in more than 600 papers (Google Scholar citation report, October 2006; click here to find the current result).

In Section 1, "Introduction," the authors briefly introduce basic concepts of statistical experiments, such as hypotheses, samples, variables, and probability distributions. In Section 2, "Estimation," the authors explain estimation process based on a sample taken from a population. The reader can learn many issues related to estimation, including estimation of confidence intervals and various estimation methods, such as maximum likelihood, ordinary

least squares, and resampling; a brief introduction to Bayesian approach is also provided. In Section 3, "Hypothesis testing," the reader is introduced to one of the main statistical concepts of the 20th century: hypothesis testing. The authors present basic concepts and explain randomization tests, multiple testing, and meta-analysis. They also present different points of view on the general idea of hypothesis testing.

In Section 4, "Graphical exploration of data," the authors teach exploratory analysis and using graphs. They also introduce the basic concepts of data transformation, standardization, outliers, and censored and missing data.

In Sections 5, "Correlation and regression," and 6, "Multiple regression and complex regression," the reader is introduced to simple and some advanced issues of one of the main statistical issues in biology: regression analysis. Some of these issues are, for instance, simple and multiple linear regression analysis, smoothing, and nonlinear models, as well as a brief introduction to structural equation modeling, which is applicable to path analysis—an important statistical tool used to understand complex causal systems.

In Section 7, "Design and power analysis," the authors introduce sampling designs—an important issue in ecological studies—and a general overview of experimental design and power analysis. This section serves as background for subsequent sections that introduce the statistical methodology that is most often applied in crop science: the analysis of variance. Section 8, "Comparing groups of treatments — analysis of variance," provides basic knowledge of ANOVA under one-way, completely randomized design experiments. The authors focus on basic ideas of ANOVA, e.g., assumptions, diagnostics, robust approaches, and multiple range tests. Section 9, "Multifactor analysis of variance," contains a description of more advanced ANOVA issues, such as analyses under multi-factorial and nested designs. In Section 10, "Randomized blocks and simple repeated measures: unreplicated two factor design," the reader learns about block designs, often exploited in crop science, and repeated measures. Repeated measures are quite often obtained in crop science investigations. Therefore, this section should help crop scientists understand this important issue and avoid mistakes. Section 11, "Split-plot and repeated measures designs: partly nested analyses of variance," is an introduction to simple and advanced issues of split-plot design—an often used design in factorial field experiments. The authors compare repeated measures designs with split-plot designs.

In section 12, "Analysis of covariance," the authors provide a thorough description of analysis of covariance (ANCOVA). The reader learns about applying ANCOVA to single-factor designs, assumptions, robust ANCOVA, and specific comparisons between means. Application of ANCOVA to more complex designs is also discussed.

In section 13, "Generalized linear models and logistics regression," general linear models, Poisson regression, general additive models, and models for correlated data are briefly introduced; logistic regression is presented more thoroughly. Section 14, "Analyzing frequencies," is an introduction to analyzing contingency tables and log-linear models.

In the next four sections, the authors discuss multivariate statistical methods, which nowadays are becoming a very important tool for crop scientists. Section 15, "Introduction to multivariate analysis," explains the importance and necessity of applying multivariate analyses in biological investigations. The reader is introduced to basic concepts of multivariate analysis, such as, among others, linear combinations, eigenvalues, and eigenvectors; distance and dissimilarity measures; multivariate outliers; and missing observations. This section shows the reader importance of, and helps them understand, specificity of multivariate statistics. In Section 16, "Multivariate analysis of variance and discriminant analysis," Section 17, "Principal components and correspondence analysis," and Section 18, "Multidimensional scaling and cluster analysis," an introduction to the most important multivariate statistical methods is presented.

Finally, in Section 19, "Presentation of results," the authors provide a valuable guide to presentation of results, including tables and graphs. They discuss some important aspects of

oral presentations, which should be useful not only to graduate students but also to professional investigators.

This book should be useful for graduate students and all crop scientists, and in general—biologists, who wish to acquire knowledge of statistics on a basic and advanced level. I believe that this book is, and will be, a key resource for those who apply statistical methods to biological data.

MARCIN KOZAK
Department of Biometry
Warsaw Agricultural University
Nowoursynowska 159, 02-776 Warsaw, Poland
m.kozak@omega.sggw.waw.pl