Biochemistry and Molecular Biology Education
Vol. 39, No. 3, pp. 242–243, 2011

# Book Reviews

**Statistics at the Bench: A Step-by-Step Handbook for Biologists**

M. Bremer and R. W. Doerge, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2010, ISBN: 978-087969857-7

## GENERAL ASSESSMENT

The authors present a very useful and a readable guide to the biological applications of statistics, and they state their purpose very clearly in their introduction: "this is a bench side manual that is designed and intended to be used by people who are in need of a quick refresher or a big-picture overview of a statistical procedure." Briefly, they achieved their purpose.

The book consistently presents and interprets the language of the statistician so that a bench scientist with limited or no formal training in statistics can grasp the principles being discussed. Each chapter presents abstract ideas and just enough theory to enable users to move on to the specific examples that are provided. Although they mention a number of dedicated statistical software packages, the authors chose to present their examples with specific instructions on data manipulation and evaluation in Microsoft Excel, "because it is one of the most commonly available software applications and is widely used by biologists today."

## CHAPTER 1 INTRODUCTION

The authors describe the increasing need for statistical evaluation as biology becomes more data intensive. Also, they explain their approach in their writing and the reasons they presented their data analysis with command scripts for Microsoft Excel.

## CHAPTER 2 COMMON PITFALLS

This chapter focuses primarily on vocabulary, explaining common mistakes or misconceptions that happen in conversations with statisticians. Some of the terminology was confusing me. My training in analytical chemistry identified two sources of error: random and systematic. In this text, the authors list biological variation and technical variation. It is unclear to me if these are different from random error (biological variation) and systematic error (technical variation).

## CHAPTER 3 DESCRIPTIVE STATISTICS

Once again, there are some very nice explanations of terminology (variable type, for example), with examples that are easily worked for the most part. One exception is on page 20, where they describe the use of the Histogram tool in Microsoft Excel. This tool is not available in Excel 2011 for Macintosh. A quick Google search revealed that a package called StatPlus LE (free download from AnalystSoft) could make these tools available. However, installation of the software was not intuitive; the web site promised a better experience if you download the commercial version ($120 educational rate). Their treatment of probability, normal distributions and binomial distributions was very well written and helpful. There is an extensive discussion of the central limit theorem, but there is no clear statement of the theorem itself.

## CHAPTER 4 DESIGN OF EXPERIMENTS

This material is very helpful, especially for scientists early in their careers. The issues of sample selection, randomization of samples, model selection and sample size are described in simple and straightforward terms. This chapter also lays the groundwork for the remaining chapters of the book; statistical terms that were introduced in chapters 1-3 are given more depth.

## CHAPTER 5 CONFIDENCE INTERVALS

The authors use a wonderful example that clearly distinguishes confidence intervals from probabilities. Section 5.1, "Interpretation of Confidence Intervals" describes the confidence level of effective birth control for a population over a one year period (98%) to the chances of pregnancy following intercourse, which must be either 0 (not pregnant) or 1 (pregnant). While presenting the issue of sample size and confidence intervals, the authors discussed t-distributions without explicitly stating the circumstances where a t-distribution would be appropriate.

## CHAPTER 6 HYPOTHESIS TESTING

One of the terms a researcher may encounter when discussing experimental design with a statistician is the "null hypothesis." This term is described in excellent detail in a variety of settings. The chapter includes a systematic approach for performing a hypothesis test and demonstrates how to use this approach with a number of different tests (three types of *t*-test, the chi-squared test and the *F*-test, among others). There is also a brief but effective explanation of parametric and nonparametric tests, with excellent examples that explain when each test is appropriate. The chapter concludes with a explanation of the E-value that is found in BLAST searches.

## CHAPTER 7 REGRESSION AND ANOVA

The authors explain clearly when to use regression and ANOVA (analysis of variance). Also, they continue to emphasize what statistics can do for you. For example, they repeatedly refer to normally distributed residuals in either type of analysis. The limitations of Excel are clearly spelled out, especially when using ANOVA models. This is particularly clear in the use of an ANOVA model for the analysis of microarray data (an issue close to the heart of many biologists and biochemists), where a four-way

ANOVA model is frequently required; Excel is limited to two-way ANOVA models.

### CHAPTER 8 SPECIAL TOPICS

This chapter contains many terms that are frequently encountered in the scientific literature. As I read this chapter, I realized that I did not really understand many of these terms (such as principle component analysis and maximum likelihood), which had seemed intuitively obvious to me. In most cases, their explanation was sufficient to bring clarity; at the very least, it pointed me in a direction that could lead to better understanding. Their explanation of clustering, particularly as it applies to phylogenetic trees, was easy to follow. Principle component analysis was explained briefly, but I was disappointed that they did not include a sample problem. Their explanation of the challenges associated with microarray data analysis (Section 8.4) was very helpful and should apply to almost all "omics" fields, which typically have large data sets, but limited numbers of replications. The section on normalization addressed a question that I frequently have as I review statistical analysis in the biochemical literature, "Why are some corrections acceptable while others are not?" The chapter ends with a comparison of frequentist (or classical) statistics and Bayesian statistics. This is one of the few sections of the book that I wish had been expanded just a bit.

### SUMMARY

This manual reads very well. It clearly fulfills its premise—to provide a useful tool to have at hand on the laboratory bench. Clearly, a 157-page book cannot provide a scientist with the background needed to perform rigorous statistical evaluation of data from the laboratory or the classroom. However, it will enable you to engage informed colleagues in meaningful discussions that will advance your research and may contribute to better success in getting grants funded or manuscripts published. The authors do a nice job of writing an internally consistent text that contains many internal references. For example, Section 7.2 on ANOVA begins with a reference to Section 3.1 to insure that the reader clearly grasps the concept of categorical predictors. For scientists who lack formal training in statistics, I recommend this book very highly and suggest creating a glossary of terms that arise frequently (variable types, common hypothesis tests, distribution types).

Paul Craig
*153 Lomb Memorial Drive, Rochester,*
*New York 14623-5608*

**DOI 10.1002/bmb.20497**

Used Books Books To Read Next Generation Sequencing Normal Distribution Standard Deviation Science Geek Biologist Statistics Reading Online. Statistics at the Bench: A Step-by-Step Handbook for Biologists. Buy Statistics at the Bench: A Step-by-Step Handbook for Biologists on Amazon.com âœ“ FREE SHIPPING on qualified orders. kristindownie science geek. Medical Laboratory Science Geek Med School Microbiology Science Classroom School Stuff Allah Geek Stuff Chain. "This book is a convenient bench-side companion for biologists, designed as a handy reference guide for elementary and intermediate statistical analyses. Statistical methods most frequently used in publications and reports, as well as guidelines for the interpretation of results, are explained Full description. Saved in Statistics at the bench by M. Bremer, 2009, Cold Spring Harbor Laboratory Press edition, in English.Â  Statistics at the bench. a step-by-step handbook for biologists. by M. Bremer. Published 2009 by Cold Spring Harbor Laboratory Press in Cold Spring Harbor, N.Y . Written in English. Subjects. Biometry. There's no description for this book yet. Can you add one? Edition Notes.