Statistical Design (Book Review)

Author:
Hibbert, D. Brynn

Publication details:
Chemometrics and Intelligent Laboratory Systems
v. 85
Chapter No. 1
pp. 155-155
0169-7439 (ISSN)

Publication Date:
2007

Publisher DOI:
http://dx.doi.org/10.1016/j.chemolab.2006.06.002

Book review

Statistical design - chemometrics
Data Handling in Science and Technology – Volume 28

R.E. Bruns, I.E Scarminio and B. De Barros Neto
Elsevier B.V., Amsterdam, 2006
422 pp, ISBN 0-444-52181-X
Hardcover EUR 130, GBP 90, USD 145

Available in Europe, Middle East and Africa from Elsevier, Customer Service Department, Linacre House, Jordan Hill, Oxford OX2 8DP, UK, and in the USA and Canada from Elsevier, Customer Service Department, 11830 Westline Industrial Drive, St. Louis, MO 63146, U.S.A.

This entertaining book is a translation from the Portuguese (*Como Fazer Experimentos*), and does not suffer at all because of that. It is a basic, minimal mathematical, introduction to experimental design, full of interesting examples, that reflect its geographical origin. Quite why the word 'chemometrics' has found its way into the title, and why 'experimental' (which I guess could be a translation of *Experimentos*) was left out I am not sure.

The first chapter is a rationale for applying statistics, starting with a quote from Descartes in 1637. It explains why the apparently common sense approach to optimization of 'change one factor at a time' is not the best, and that statistics can help. The goal of experimental design is to extract the maximum amount of information from a minimum outlay of experiments. Some fundamental concepts concerned with the requirements to describe systems with empirical models versus explaining systems using theoretical models, are discussed. This allows the introduction of experimental design and the different kinds of questions that can be asked about a system, to end the first chapter.

The obligatory rehearsal of statistical principals comes in chapter 2. The concept of random variation of results is introduced by an example of acetic acid titrations, with "exaggerations for didactical purposes". It must be said that the data looks too random and has no outliers. (The author of this review has just published a paper on the use of data from acetic acid titrations – a schools' titration competition which is rich with real statistical fodder [1].) Examples in this chapter include the masses of beans: black beans, *roxinho* beans which are purple, and *carioca* beans whose colour we do not learn. The footnotes should be read as asides to ease the strain of the text, such as "life is to short anyway" or "or he, et cetera" or "also for want of a better idea". This chapter is a good general introduction of frequentist statistics and covers what is needed for the rest of the book.

Chapter 3 –*Changing everything at the same time*, is the introduction to factorial designs. How they are set up, the analysis to find main and interaction effects and normal probability plots to determine significant effects are covered. It would have been useful to show the Rankit procedure to create a normal probability graph, as from the explanation here it is not quite obvious what the analyst actually does. Fractional designs, including Taguchi and Plackett-Burman designs are treated in

Chapter 4. The resolution of a fractional design is explained well, and the nine examples cover the concepts in the chapter.

Empirical model building is introduced in Chapter 5. From the linear model to more complex systems, the assessment of how well the model fits the data is explained. How to conduct an analysis of variance and calculate confidence intervals on the parameters are described using matrix arithmetic. The first example is a nice re-working of Robert Boyles' 1661 experiments on the pressure volume relations of air. Not surprisingly a linear model does not do at all well, but a cubic function fits the PV data with 99.8% of the variance explained. Boyle's law ($P \propto 1/V$) fits best of all.

Response surface methodology (Chapter 6), styled RSM, describes the combination of modeling and steepest ascent optimization. Derringer and Suich desirability criteria are recommended to monitor the path to the optimum. Designs described include Central Composite, Box-Behnken, Doehlert and alphabetic-Optimal. Mixture models are treated in chapter 7 and the book finishes with Simplex optimization.

The authors stress that they are chemists and not statisticians and posit that readers of the book need the minimum math and stats to achieve their chemical ends. I think they have succeeded. The examples are good and answers to exercises are included at the end of the book. I recommend this book for university libraries and as a source book for educators and students.

Reference

[1] Hibbert, Brynn *Teaching modern data analysis with the Royal Australian Chemical Institute's titration competition* Aust. J. Ed. Chem. **66**, 5-11 (2006).

Professor D Brynn Hibbert
School of Chemistry
University of New South Wales
Sydney, Australia

Contact:
Telephone: +61 2 9385 4713
Fax: +61 2 9385 6141
e-mail: b.hibbert@unsw.edu.au