Algorithmic Operations Research

Review of the electronic book “Dealing with Uncertainties”

Michael Morten Steurer

Volume 3, numéro 1, winter 2008

URI : https://id.erudit.org/iderudit/aor3_1br02

Aller au sommaire du numéro

Éditeur(s)

Preeminent Academic Facets Inc.

ISSN

1718-3235 (numérique)

Découvrir la revue

Résumé de l'article

I review the recent book authored by Manfred Drosg, Dealing with Uncertainties, Springer, Berlin Heidelberg New York 2007, ISBN-10 3-3-540-29606-9, ISBN-13 978-3-540-29606-5, which is the English translation of the author’s German book, Der Umgang mit Unsicherheiten, Facultas Universit¨atsverlag, 2006, ISBN 3-85076-748-5

Algorithmic Operations Research Vol.3 (2008) 94–96

Book Reviews

Review of the electronic book "Dealing with Uncertainties"

Michael Morten Steurer

Department for Neurobiology and Cognition Research, University of Vienna, Italy

Abstract

I review the recent book authored by Manfred Drosg, Dealing with Uncertainties*, Springer, Berlin Heidelberg New York 2007, ISBN-10 3-3-540-29606-9, ISBN-13 978-3-540-29606-5, which is the English translation of the author's German book,* Der Umgang mit Unsicherheiten*, Facultas Universitätsverlag, 2006, ISBN 3-85076-748-5.*

1. Introduction

Uncertainty is a key feature of modern science. Acknowledging/accepting the uncertainty of knowledge, its measurement, the operational dealing with it as well as the publication of the latter are essential parts of scientific work. Carl Sagan writes in his bestseller *The Demon-Haunted World* [1]:

> Science is far from a perfect instrument of knowledge. It's just the best we have.

Some paragraphs later he writes:

> Every time a scientific paper presents a bit of data, it's accompanied by an error bar—a quiet but insistent reminder that no knowledge is complete or perfect. It's a calibration of how much we trust what we think we know. If the error bars are small, the accuracy of our empirical knowledge is high; if the error bars are large, then so is the uncertainty in our knowledge. Except in pure mathematics, nothing is known for certain (although much is certainly false).

Uncertainty does not origin from sloppy or faulty measurements, but is rather a characteristic of the scientific context in which the data is embedded. It is an intrinsic quality of the underlying system, resp. the scientific model that is used to interpret the data or the process of measuring. Faulty measurements lead to deviations which have to be corrected, but do not lead to uncertainty. The magnitude and the sign of these deviations can be calculated (e.g. the load of an electronic circuit by using a voltmeter) and thus corrected, even if the deviation itself cannot be prevented. By contrast, the actual magnitude and the sign of effects that are associated with uncertainty cannot be calculated.

Although uncertainty can be minimized, its incidence is unavoidable. Deviation and uncertainty are completely different matters. Although measuring often inevitably leads to uncertain results, it is not the cause of all errors. Uncertainties stem from at least two sources: First, the use of stochastic models of reality, which are either statistical models of reality (e.g. statistical mechanics, thermodynamics) or probabilistic models (e.g. radioactive decay, quantum mechanics). Second, the (unavoidable) mathematical incommensurability of certain quantities (Hippasos of Metapont). By using floating point numbers the consequences of the latter are evident. Irrational as well as most of the rational numbers can not be represented exactly. Generally, the represented number deviates from the *true* value, which then is inaccessible. The magnitude and the sign of this deviation will always remain unknown, although its boundaries can be specified. The *digitizing uncertainty* is a probabilistic measure of that matter.

In addition to characterizing the above described effects uncertainty opens up the possibility to select among multiple competing theories. The principle of *Occam's razor* states that the explanation of any phenomenon should make as few assumptions as possible by eliminating those that make no difference in the observable predictions. Uncertainties are the attributes that make it possible to decide whether differences do or do not exist. Even more important, dealing with uncertainties offers the possibility to judge whether or not data is overfitted.

Whenever information is framed in a scientific scope it becomes two-dimensional: data and its uncertainty. Uncertainty is not the scientific stopgap of incapacity in describing nature (or the cosmos). In fact it is an essential principle of modern science.

Several books about *error analysis* have been written. Very often the distinction between deviation and uncertainty is neglected. This may have to do with the use of the very broad terminus *error*. Also, the application of probability theory and statistics to the phenomenon of incomplete knowledge may have contributed to the widespread habit of assessing uncertainty inductively, i.e., by calculating the standard error from the fluctuation of the data. According to this approach, correlations between components of uncertainty cannot be assessed (e.g. the scale uncertainty can never be induced by repeated measurements), and, even worse, it is impossible to determine the uncertainty of a single data value. As a young student of physics I was often puzzled and really became frustrated when trying to understand the multitude of inconsistent approaches to uncertainty in the scientific literature. The present book by Manfred Drosg offers a consistent way of dealing with uncertainty.

In particular, the book contributes to the understanding of and dealing with uncertainty in the following ways:

- Uncertainty is understood as a principle of data embedded in a scientific context rather than as a stopgap for sloppy measurement. The book addresses the question of how to deal with data instead of just focusing on how to handle measurement outcomes.
- The distinction between deviation and uncertainty is clearly made and the misleading terminus *error* is avoided.
- To the best of my knowledge this is the first book that presents the classical inductive approach to external uncertainty (precision) as well as the deductive approach to internal uncertainty (accuracy). It is shown that the relation between the two types of uncertainty can be used as a basis upon which to decide whether the theoretical model used to describe the data is overdetermined (overfitted) or underdetermined. Also, the book touches on the application of the deductive approach in risk estimation.
- The main sources of uncertainty in measurement and the effects of data uncertainty on distributions are described in a general way. The distortion effects of differential and integral nonlinearity in the conversion of analog to digital data values are explained.
- Feedback of uncertainties on experiment design and time as well as cost optimization of experiments by minimizing uncertainty are discussed.
- The estimation of uncertainties as well as the question of how to calculate with them are comprehensively discussed by means of many examples. The appropriate presentation of scientific data and their uncertainties is outlined.

2. Contents

Interestingly, the book starts with a prolog in which seven *myths in error analysis* are enumerated (e.g. Myth 7: *It is all right to "guess" an error.*) and their falsity is revealed. This prolog raises the readers' attention and makes them eager to go on reading. As the book title already tells us, the author mainly addresses the operational dealing with uncertainty rather than philosophical and abstract theoretical concepts of the matter. Only the first chapter touches on the philosophical and theoretical aspects. The second chapter deals with the basics on data. The question "What is a measurement?" as well as the differences between analog and digital quantities are addressed. Furthermore, truncation and unbiased consecutive rounding of numbers are discussed.

The third chapter illustrates the basics on uncertainty and introduces the relevant definitions. The three main sources of uncertainty of the measuring process are explained, namely the *digitizing uncertainty, scale uncertainty and interpolation uncertainty (nonlinearity)* and uncertainties stemming from rounding are discussed. The correct quotation of uncertainties, the properties of tolerances (maximum uncertainties) as well as the quadratic summation of uncertainties and its properties are outlined. Furthermore the topic of outliers is touched and the options of dealing with them are enumerated.

Chapter 4 introduces the radioactive decay as a model for random events. Event counting and the problem of counting loss due to dead time are discussed and the relevant corrections are explained. The model of radioactive decay is then used to introduce the inductive approach to uncertainty. The chapter closes with properties of data sets, measures of dispersion, reproducibility and an introduction into linear regression.

In Chapter 5 frequency and probability distributions and their characteristics are described. The effects of data uncertainty—especially the integral and differential nonlinearity—on frequency distributions are discussed. Dealing with probabilities, stochastic independence and statistical confidence are addressed.

Chapters 6 to 8 are mainly dedicated to the deductive approach to uncertainty. The principle of deducing uncertainty from internal features of the quantity as well as from the measuring process is introduced and the relations between inductively assessed external uncertainties and the deduced internal uncertainties are outlined. The topic of regression is discussed in detail and directs the reader's attention to the problem of data consistency and correlation. The *general law of error propagation* and the correct calculation of total uncertainty consisting of multiple uncorrelated as well as correlated components are explained. Many examples safely guide the novice through the complex material.

The appropriate presentation of data and uncertainties is comprehensively covered in chapter 9. It includes a discussion of optimal scaling of the coordinates as well as zero point suppression in diagrams and the use of uncertainty bars or uncertainty rectangles.

The 10th and last chapter is entirely dedicated to the feedback of uncertainties on experiment design. Time and cost optimization of experiments by minimizing uncertainty is discussed and illustrated by means of many examples. The book ends with an impressive example of the *Ratio Method*, by which the total uncertainty is minimized through mutual compensation of systematic (i.e., correlated) uncertainty components.

3. Conclusion

By means of many examples the author tries to provide the reader with an intuitive, practical account to uncertainty. Although the book mainly addresses physicists and students of physics, it is, in my opinion, of value also to other scientists. Chemists, biologists and computer scientists may find the concept of uncertainty very useful in their work.

As an example I may spend a few lines on uncertainty's application in computer science. Numerical computation usually deals with maximum errors only (see for example [2]). The concept of uncertainty may have some impact on that field because it opens more options than the concept of maximum error does (e.g. applying *Occam's razor*). The concepts of differential and integral nonlinearity may contribute to the work of computer scientists dealing with resolution problems in imaging. Although much effort has been made in this field (see for example [3]), to the best of my knowledge, the frequency distorting effects of nonlinearity in the measuring process are, for the most part, modeled as uncorrelated noise only, instead of interpolation uncertainty. Although the author only shows the 1-dimensional case, its principle can be adapted to more-dimensional frequency distributions.

I explicitly recommend this book to all scientists and students that find the usual (in my opinion inconsistent) descriptions and compositions of many books and courses dissatisfying and frustrating, as it is written clearly, consistently, conclusively and close to practice.

Received 10 January, 2007

References

[1] Carl Sagan, 1996: *The Demon-Hounted World, Science as a Candle in the Dark*, A Ballantine Book, Published by The Random House Publishing Group. ISBN 0-345-40946-9.

[2] Christoph W. Ueberhuber, 1997: *Numerical Computation: methods, software, and analysis*, Springer, Berlin Heidelberg New York ISBN 3-540-62058-3.

[3] Marteza Shahram, and Peyman Milanfar: *Statistical and Information-Theoretic Analysis of Resolution in Imaging*, IEEE Transactions on Information Theory, Vol. 52 No. 8, August 2006