

Preface

This book is written as a guide for the presentation of experimental data including a consistent treatment of experimental errors and inaccuracies. It is meant for experimentalists in physics, astronomy, chemistry, life sciences and engineering. However, it can be equally useful for theoreticians who produce simulation data: they are often confronted with statistical data analysis for which the same methods apply as for the analysis of experimental data. The emphasis in this book is on the determination of best estimates for the values and inaccuracies of parameters in a theory, given experimental data. This is the problem area encountered by most physical scientists and engineers. The problem area of experimental design and hypothesis testing – excellently covered by many textbooks – is only touched on but not treated in this book.

The text can be used in education on error analysis, either in conjunction with experimental classes or in separate courses on data analysis and presentation. It is written in such a way – by including examples and exercises – that most students will be able to acquire the necessary knowledge from self study as well. The book is also meant to be kept for later reference in practical applications. For this purpose a set of “data sheets” and a number of useful computer programs are included.

This book consists of parts. Part I contains the main body of the text. It treats the most common statistical distributions for experimental errors and emphasizes the error processing needed to arrive at a correct evaluation of the accuracy of a reported result. It also pays attention to the correct reporting of physical data with their units. The last chapter considers the inference of knowledge from data from a Bayesian point of view, hopefully inducing the reader to sit back and think. The material in Part I is kept practical, without much discussion of the theoretical background on which the various types of analysis are based. This will not at all satisfy the eager student who has sufficient background in mathematics and who wishes to grasp a fuller understanding of the principles involved. Part II is to satisfy the curious: it contains several Appendices that explain various issues in more detail and provide derivations of the equations quoted in Part I. The Appendices in

Part II obviously require more mathematical skills (in particular in the field of linear algebra) than Part I. Part III contains Python code examples and Part IV provides answers to exercises. Finally, Part V contains practical information in the form of “data sheets” which provide reference data in a compact form.

Throughout the book computer programs are included to facilitate the computations needed for applications. There are several professional software packages available for statistical data analysis. In the context of an educational effort, I strongly advise against the use of a specialized “black-box” software package that can be easily misused to produce ill-understood results. A “black-box” computer program should never be a magic substitute for a method that is not understood by the user! If a software package is to be used, it should provide general mathematical and graphical tools, preferably in an interactive way using an interpreter rather than a compiler. The commercial packages MATHEMATICA, MATLAB and MATHCAD are suitable for this purpose. However, most readers of this book will not have access to any or all of these packages, or – if they have temporary access through their institution – may not be able to continue access at a later point in time. Therefore for this book the choice was made to use the generally available, actively developing, open-source interpretative language PYTHON. With its array-handling and scientific extensions NUMPY and SCIPY the capabilities of this language come close to those of the commercial packages. Software related to this book, including a Python module `plotsvg.py` providing easy plotting routines, can be found on www.hjcb.nl/.

This book is the successor of the Dutch textbook *Goed meten met fouten* (Berendsen, 1997) that has been used in courses at the departments of physics and chemistry of the University of Groningen since 1997. The author is indebted to Emile Apol, A. van der Pol and Ruud Scheek for corrections and suggestions. Comments from readers are welcome to author@hjcb.nl.