

# Preface

The more questions you ask, the more wrong answers you are expected to receive—even if every single source of your information is quite trustworthy. In this work, the sources of information are data, and the questions are formalized by statistical hypothesis-alternative pairs. From the mathematical point of view, this leads to multiple test problems. We will discuss criteria and methods (in particular multiple tests) which ensure that with high probability not too many wrong decisions are made, even if many hypotheses are of interest under the scope of one and the same statistical model, i.e., regarding one and the same dataset.

High-throughput technologies in different fields of modern life sciences have led to massive multiplicity and given rise to multiple test problems with more hypotheses than observations. Driven by these developments, also new statistical paradigms have arisen. It is fair to say that a new era of multiple testing began when Yoav Benjamini and Yosef Hochberg formally introduced the false discovery rate (FDR) and the linear step-up test for FDR control in 1995. In this book, apart from classical methods controlling the family-wise error rate (FWER), theory and important life science applications of the FDR are presented in a systematic way, presumably for the first time in this depth in a monograph. In this, focus is on frequentist approaches aiming at FDR control at a fixed level. Other type I and type II error rates are mentioned and discussed where appropriate, but focus is on FWER and FDR. [Chapters 6 and 7](#) broaden the view and show how multiple testing methodology can be used in the context of binary classification and model selection, respectively, with life science applications provided in Parts II and III. Further relationships between multiple testing and other simultaneous statistical inference problems are discussed in [Chap. 1](#) and at respective occasions.

The book is primarily meant to be a research monograph and an introduction to simultaneous inference for applied statisticians and practitioners from the life sciences. To this end, presentation is with emphasis on applicability and we provide a couple of hints concerning which multiple test to use for which type of data. Furthermore, [Chap. 8](#) deals with software implementing the theoretically treated procedures. However, the mainly theoretical Part I of the book may also serve as the basis for a graduate course on simultaneous statistical inference with emphasis on multiple testing for mathematical statisticians. I used parts of [Chaps. 2, 4 and 5](#) for such a course at Humboldt-University Berlin and a couple of diploma theses in mathematics originated from this teaching.

The material for this book originated from joint work with many colleagues. I acknowledge the respective co-workers at the end of each chapter. Apart from his scientific contributions, I am especially grateful to Taras Bodnar who critically read every chapter and provided many constructive comments which helped to improve the presentation.

My deepest gratitude, however, is due to the Thai branch of my family for their enduring support in hard times. Therefore, I dedicate this work to Prayun, Duangchan, Dako, Pipat, Janyarak and her children, and my wife Supansa.

Berlin, January 2014

Thorsten Dickhaus

Simultaneous Statistical Inference  
With Applications in the Life Sciences

Dickhaus, T.

2014, XIV, 180 p. 19 illus., 10 illus. in color., Hardcover

ISBN: 978-3-642-45181-2