

Prefaces to the 1st edition

Part I

This small book addresses different kinds of data files, as commonly encountered in clinical research and their data analysis on SPSS software. Some 15 years ago serious statistical analyses were conducted by specialist statisticians using main-frame computers. Nowadays, there is ready access to statistical computing using personal computers or laptops, and this practice has changed boundaries between basic statistical methods that can be conveniently carried out on a pocket calculator and more advanced statistical methods that can only be executed on a computer. Clinical researchers currently perform basic statistics without professional help from a statistician, including t-tests and chi-square tests. With the help of user-friendly software, the step from such basic tests to more complex tests has become smaller and more easy to take.

It is our experience as masters' and doctorate class teachers of the European College of Pharmaceutical Medicine (EC Socrates Project, Lyon, France) that students are eager to master adequate command of statistical software for that purpose. However, doing so, albeit easy, it still takes 20–50 steps from logging in to the final result, and all of these steps have to be learned in order for the procedures to be successful.

The current book has been made intentionally small, avoiding theoretical discussions and highlighting technical details. This means that this book is unable to explain how certain steps were made and why certain conclusions were drawn. For that purpose additional study is required, and we recommend that the textbook "Statistics Applied to Clinical Trials," Springer 2009, Dordrecht, Netherlands, by the same authors, be used for that purpose, because the current text is much complementary to the text of the textbook.

We have to emphasize that automated data analysis carries a major risk of fallacies. Computers cannot think and can only execute commands as given. As an example, regression analysis usually applies independent and dependent

variables, often interpreted as causal factors and outcome factors. For example, gender or age may determine the type of operation or type of surgeon. The type of surgeon does not determine the age and gender. Yet a software program does not have difficulty to use nonsense determinants, and the investigator in charge of the analysis has to decide what is caused by what, because a computer cannot do things like that, although they are essential to the analysis. The same is basically true with any statistical tests assessing the effects of causal factors on health outcomes.

At the completion of each test as described in this book, a brief clinical interpretation of the main results is given in order to compensate for the abundance of technical information. The actual calculations made by the software are not always required for understanding the test, but some understanding may be helpful and can also be found in the above textbook. We hope that the current book is small enough for those not fond on statistics but fond on statistically proven hard data in order to *start on SPSS*, a software program with an excellent state of the art for clinical data analysis. Moreover, it is very satisfying to prove from your own data that your own prior hypothesis was true, and it is even more satisfying if you are able to produce the very proof yourself.

Lyon, France
December 2009

Ton J. Cleophas
Aeilko H. Zwinderman

Part II

The small book “SPSS for Starters” issued in 2010 presented 20 chapters of cookbook-like step by step data analyses of clinical research and was written to help clinical investigators and medical students analyze their data without the help of a statistician. The book served its purpose well enough, since 13,000 electronic reprints were being ordered within 9 months of the edition.

The above book reviewed, e.g., methods for:

1. Continuous data, like t-tests, nonparametric tests, and analysis of variance
2. Binary data, like crosstabs, McNemar’s tests, and odds ratio tests
3. Regression data
4. Trend testing
5. Clustered data
6. Diagnostic test validation

The current book is a logical continuation and adds further methods fundamental to clinical data analysis.

It contains, e.g., methods for:

1. Multistage analyses
2. Multivariate analyses
3. Missing data

4. Imperfect and distribution free data
5. Comparing validities of different diagnostic tests
6. More complex regression models

Although a wealth of computationally intensive statistical methods is currently available, the authors have taken special care to stick to relatively simple methods, because they often provide the best power and fewest type I errors and are adequate to answer most clinical research questions.

It is time for clinicians not to get nervous anymore with statistics and not to leave their data anymore to statisticians running them through SAS or SPSS to see if significances can be found. This is called data dredging. Statistics can do more for you than produce a host of irrelevant p-values. It is a discipline at the interface of biology and mathematics: mathematics is used to answer sound biological hypotheses. We do hope that “SPSS for Starters 1 and 2” will benefit this process.

Two other publications from the same authors entitled *Statistical Analysis of Clinical Data on a Pocket Calculator 1 and 2* are rather complementary to the above books and provide a more basic approach and better understanding of the arithmetic.

Lyon, France
January 2012

Ton J. Cleophas
Aeilko H. Zwinderman

Preface to 2nd edition

Over 100,000 copies of various chapters of the first edition of SPSS for Starters (Parts I (2010) and II (2012)) have been sold, and many readers have commented and given their recommendations for improvements.

In this 2nd edition, all the chapters have been corrected for textual and arithmetic errors, and they contain updated versions of the background information, scientific question information, examples, and conclusions sections. In “notes section”, updated references helpful to a better understanding of the brief descriptions in the current text are given.

Instead of the, previously published, two-20-chapter Springer briefs, one for simple and one for complex data, this 2nd edition is produced as a single 60-chapter textbook.

The, previously used, rather arbitrary classification has been replaced with three parts, according to the most basic differences in data file characteristics:

1. Continuous outcome data (36 chapters)
2. Binary outcome data (18 chapters)
3. Survival and longitudinal data (6 chapters)

The latter classification should be helpful to investigators for choosing the appropriate class of methods for their data.

Each chapter now starts with a schematic overview of the statistical model to be reviewed, including types of data (mainly continuous or binary (yes, no)) and types of variables (mainly outcome and predictor variables).

Entire data tables of the examples are available through the Internet and are redundant to the current text. Therefore, the first 10 rows of each data table have now been printed only.

However, relevant details about the data have been inserted for improved readability.

Also simple explanatory graphs of the principles of the various methods applied have been added.

Twenty novel chapters with methods, particularly, important to clinical research and health care were still missing in the previous edition, and have been added.

The current edition focuses on the needs of clinical investigators and other nonmathematical health professionals, particularly those needs, as expressed by the commenters on the first edition.

The arithmetic is still more of a no-more-than high-school level, than that of the first edition, while complex computations are described in an explanatory way.

With the help of several new hypothesized and real data examples, the current book takes care to provide step-by-step data-analyses of the different statistical methodologies with improved precision.

Finally, because of lack of time of this busy group of people, as expressed by some readers, we have given additional efforts to produce a text as succinct as possible, with chapters, sometimes, no longer than three pages, each of which can be studied without the need to consult others.

Lyon, France
January 2015

Ton J. Cleophas
Aeilko H. Zwinderman

<http://www.springer.com/978-3-319-20599-1>

SPSS for Starters and 2nd Levelers

Cleophas, T.J.; Zwinderman, A.H.

2016, XXV, 375 p. 148 illus., 30 illus. in color.,

Hardcover

ISBN: 978-3-319-20599-1