

# **Statistics Companion**

**The Essential Guide to Statistical Textbooks**

Including Excel Guidance



Erik M. Bøye

# STATISTICS Companion

**The Essential Guide to Statistical Textbooks**

Including Excel Guidance

*Swismark*

## **STATISTICS COMPANION**

- **The Essential Guide to Statistical Textbooks**
- Including Excel Guidance

© 2010 by Erik M. Bøye and Swismark. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of the publishing firm Swismark or Copy-Dan, Denmark.

Published by the publishing firm Swismark.  
Printed by Clemensstrykkeriet, Denmark

Printed in Denmark 2010  
ISBN 978-87-993632-2-3

*Swismark*

Rugbakkevej 12  
DK-8200 Aarhus N  
Denmark  
+45 86 10 16 41

Website address:  
[www.swismark.dk](http://www.swismark.dk)  
Email address:  
swismark@swismark.dk

# Preface

This little book is a *helping hand* to users of textbooks in statistics, which often, in relation to shape and size, resemble nothing more than a brick. This can be off-putting and for way too many students, their textbook in statistics literally remains a closed book.

During my teaching of statistics at the Aarhus School of Business and at the University of Southern Denmark I have received countless enquiries from frustrated students who got stuck with their textbook. Therefore, in 2004, I published my edited lecture notes and slides. The many positive responses I received from relieved students indicate that I may have hit the nail on the head.

The current book is a translation of the second edition of *Statistik-hjælperen*. I am especially grateful to IB student Adam Blatchley Hansen, for his corrections of my translation from Danish to English.

*Statistics Companion – The Essential Guide to Statistical Textbooks* is, as mentioned above, a helping hand to understanding “bricks”. It is not strictly speaking a textbook as it leaves out many of the less important things in order to keep the book from swelling up too much. The assumptions for use of the various probability distributions are often just listed without further discussion as these assumptions are dealt with in detail in the textbooks which the *Statistics Companion* is intended to support.

The most important concepts and subjects in theoretical statistics are explained in this book. Practical statistics or descriptive statistics – which is the basis for theoretical statistics – is briefly covered at the start of the book. Full-scale insights into the tools used for descriptive statistics are given in my book *Descriptive Statistics* (2010).

The *Statistics Companion* gives detailed instructions on how to use the spreadsheet Excel (version 2007) for statistics. Elementary previous knowledge in Excel is assumed.

The statistics to be used in the exercises can be downloaded from the internet via [www.swismark.dk](http://www.swismark.dk)

Aarhus/Odense, September 2010

Erik M. Bøye

# Contents

1 Statistical Method . . . . .	1
2 Measures of Location and Measures of Dispersion . . . . .	9
3 Probability Calculation . . . . .	23
4 Discrete Probability Distributions . . . . .	33
5 The Normal Distribution . . . . .	43
6 Sampling Distribution of the Sample Mean . . . . .	55
7 Confidence Intervals . . . . .	63
8 The $\chi^2$ -distribution. Confidence Interval for $\sigma^2$ . . . . .	71
9 Hypothesis Testing . . . . .	73
10 Comparison of Two Variances: F-distribution . . . . .	81
11 Comparison of Two Means etc. . . . .	85
12 Analysis of Variance . . . . .	95
13 Simple Linear Regression and Correlation . . . . .	105
14 Multiple Regression . . . . .	121
15 $\chi^2$ -tests . . . . .	129
Annex. Statistical Tables . . . . .	135
Index . . . . .	155

## Abbreviations

mio.	million
bio.	billion = 1 000 000 000
cf.	compare with/to

## Regarding decimal separator

The Excel commands referred to in this guidance use a *point* as the decimal separator. So  $12\frac{1}{4}$  is written as 12.25.

But your computer may demand a *comma* instead. It depends on the settings of your computer.

# 1 Statistical Method

## 1.1 Sample versus population

The *statistical methods* can be divided up into the following four:

1. Methods for collecting statistics.
2. Methods for making a theoretical concept measurable.
3. Methods for handling, summarizing and presenting statistics.
4. Methods for drawing conclusion about a population from a sample.

Point 4, statistical inference, is the essence of the discipline *theoretical statistics* and the primary subject of this book.

•	Statistical inference	•
•	Methods for drawing conclusion about a POPULATION from information in a SAMPLE	•

Point 3 is treated in the discipline *descriptive statistics* or *practical statistics*. Descriptive statistics is the basis for theoretical statistics. Here we find the definitions of the two most often used concepts in textbooks in statistics – besides from *sample* and *population* – i.e.:

*Mean* or average – a measure of position or location.

*Standard deviation* – a measure of dispersion.

Chapter 2 will give a brief overview of the most important concepts of descriptive statistics.

The population depends on the survey. It may be the total population in a country or just voters, school children or pensioners. Agricultural statistics consist of several different populations: farmers, skilled farm-workers, livestock, cultivated area, farms, combines etc. The population may also be the commodities produced in an enterprise or the throw of two dice. In probability calculations the latter is an example of an infinite population.

A population consist of *units* or *elements*: i.e. the individual inhabitants, voters, farms, throws, products etc. Table 2.2 on page 21 shows a sample consisting of 36 units (houses).