test

Ovande Furtado, Jr., Ph.D.

2020-03-22

Contents

Preface

This is a companion book[1] used in a course[2] I teach at Cal State Northridge. This book is comprised of step-by-step tutorials showing how to perform analysis with SPSS and *R*.

The emphasis of this book is not to teach students how to program in *R* or SPSS. Rather, the focus is on `data analysis and interpretation` using *R* and SPSS.

Future versions of this book will contain topics related to **hypothesis testing**, **descriptive and inferential statistics**, etc. The goal is for this book to evolve into a textbook that may be used in Introductory Statistics courses taught at the undergraduate or graduate levels.

Updates

I announce major updates made to this book on Twitter. For updates, follow me at ofurtado[3]. Alternatively, you can subscribe to my blog www.drfurtado.us for updates.

How to use this book

To take full advantage of this book, I encourage you to download and install either RStudio or SPSS[4]. so that you can follow along. You will need to take some time to practice the tutorials using RStudio or SPSS. Besides, make sure to complete the challenge exercises you will encounter while studying the tutorials found throughout this book.

[1] The online version of this book is free to read and licenced under Creative Commons Attribution-NonCommercial 4.0 International License.

[2] KIN610 - Quantitative Analysis of Research in Kinesiology

[3] http://twitter.com/ofurtado

[4] *R* is free to download and use whereas SPSS is a paid application. However, there is a 14-day free trial to use SPSS

0.1 R Markdown Reference Guide

For those interested in

Chapter 1

Statistical Software

Since this ebook is being created with Biomedical students in mind, I decided cover *R* and SPSS. Wikipedia[1] has a comprehensive list of statistical software, whcih is worth checking out.

Below, I will show you how to download and install *R*, RStudio, and SPSS. I strongly recommend you to download the application of your choice and follow along.

1.1 Package Download & Instalation

In this section, you'll find all you need to get started with **RStudio** and **SPSS**. Both computer packages are available for Windows, Mac, and Linux.

1.1.1 *R* and RStudio

R is an open-source computer package used for statistical data analysis and graphics. RStudio combines several tools in one single computer package, which helps one be more productive when using *R*. Although you can use *R* by itself, the tutorials for this book were developed using RStudio. Thus, to follow along, please download both *R* and RStudio.

Standalone Version

- First, download and install *R* from this link[2]: http://bit.ly/3bBLKHO
- Then, download and install RStudio from this link: http://bit.ly/2D8H7EU

[1] List of statistical packages: http://bit.ly/2pNLjqU
[2] RStudio will detect *R* the first time you run it.

To take full advantage of *R*, you will need to install a few *R* packages[3] as they are required to run most of the statistical analysis and to create graphs.

> YouTube[4] has several video tutorials showing how to install these applications.

RStudio Cloud

If you don't want to download/install *R* and RStudio, you can use the cloud version of RStudio, provided you have access to the Internet. Simply, follow the link below to create a free account and use RStudio online, without the need to install anything in your computer.

- https://rstudio.cloud

1.1.2 SPSS

Contrary to *R*, SPSS is not free. However, you can download and use SPSS for 14 days by visiting this link: http://ibm.co/2brfyu1.

Hearne Software offers SPSS licenses that last 6, 12, or 24 months. This is specially helpful if you are taking an undergraduate or graduate course and needs access to SPSS for the time the course is being taught. Ensure to ask your instructor which package will be necessary, provided the requirements of the course. They offers *premium*, *standard*, and *base* license packages. Follow this link for more information: http://bit.ly/2w5JcBs.

1.2 Datasets

The datasets for this ebook are stored on Github[5]. If necessary, I will direct you to Github to download datasets when covering the topics for this ebook.

[3] Follow this link to learn how to install packages in *R*: http://bit.ly/2VlJ4sa
[4] Visit this link to watch a video tutorial: http://bit.ly/2OAnETa
[5] Link to Github: http://bit.ly/2VBL7bR

Chapter 2

Fundamental Concepts

2.1 Measurement

The reason is simple. Statistics would not exist without the concept of measurement. **Measurement** "is the process of comparing a value to a standard"(vincent). When Iverson measures his 3rd-grade students' proficiency on the vertical jump, he is comparing the performance with a standard, which in this particular case is how far up students can reach (inches). A measuring tape (the instrument) is used to collect bits of information. The information collected through this process is called data. Data can be quantitative or qualitative in nature. In the example above, the resulting data are quantitative, owing to the level of precision involved in the collection of data. A score of 15 inches assigned to a performer on the vertical jump test is not open for interpretation; it's a precise (objective) measurement. On the other hand, if Iverson decides to assess (not measure) the qualitative aspects of the vertical jump (the form of the movement), he would have to rely on a different - more subjective - type of instrument. The FG-COMPASS (**?**) is an example of such a test. Participants are assessed on a scale from 1 through 4, with 1 being less proficient and 4 being more proficient. Iverson and his assistant may disagree when assigning scores to a given performance. Some disagreement is accepted, even expected, between observers, owing to the subjective nature of the scale being used. Inter-rater reliability, sometimes called objectivity, can be used to verify the level of agreement.

2.2 Variables

In the context of Statistics, variables can be classified into two general types: numerical or categorical. Numerical variables can take on

2.2.1 Types of variables

2.2.2 Classification of data

2.2.3 Choosing the correct test

Choosing the correct test for a particular research study is not an easy task. This is mainly because one… book, and any statistics resource for that matter, will not be very helpful if you didn't do your homework and by that I mean:

spent time drafting your research question(s) and hypothesis(es); list your dependent variable(s) and independent variable(s) for each of the proposed hypotheses. An example will help illustrate the point being made here.To get started, let's import a dataset (data01), which is saved in my computer.

Step 1 In Step 1, we will import the data from its online location^[If working off-line, ensure to download the data prior to running .

2.3 Symbols

Below are some of the symbols I will use throughout this book to refer to various statistical terms.

	Symbol Meaning	Reference to a value in a
n	Sample size	Sample
$\bar{x}$	Sample mean	Sample
μ	Population mean	Population
s	Sample standard deviation	Sample
σ	Population standard deviation	Population
r	Pearson correlation	Sample
ρ	Pearson correlation	Population
x, z	Observed data value	Population

2.4 Equations

Although, in general, you will not be required to perform calculations by hand while following the tutorials found in this book, it is worth getting familiar with some key equations[1]. I will list them below.

Example… see Equation (**??**).

Standard error of the mean

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{2.1}$$

[1] Adapted from Stoettner (2017) (http://bit.ly/2PVJKRu)