TABLE OF CONTENTS

MOTIVATION

The most straightforward way to explain why this book exists is via a brief description of my journey from a statistical know-nothing to an author of a book on statistics.

Some years back, I set out to learn more about the application of state-of-the-art scientific methods to the business world of data-driven decision-making. Starting with analyses of observational data, I quickly shifted my focus to online controlled experiments. These are commonly referred to as A/B tests, or split tests.

At the time, I had no formal training in statistics, and only college-level understanding of mathematics, so, to be honest, I didn't even know where to start! Available books on A/B testing barely had anything to say about statistics, so I started reading online blog posts and educational resources from universities, such as online lectures and courses, as well as the odd scientific paper.

In doing so, I had to face an entirely **new jargon** full of counterintuitive terms, such as 'statistical significance', which has little to do with significance, 'statistical power' which has nothing to do with power in the casual sense, 'confidence interval', which has nothing to do with any kind of confidence, and so on. And, above all else, I had to familiarize myself with a notation full of small and capital Greek letters (α, β, γ, δ, θ, μ etc.), which would sometimes mean different things in different contexts, while different letters would also denote the same concept.

To make things worse, there were ample examples of **vague or conflicting information**. There were dozens of definitions for what a p-value is and how it should be interpreted. Almost nobody seemed to care to define what a family of hypotheses is supposed to be when discussing the Family-Wise Error Rate. One source would claim one-tailed tests are preferable, while others would swear by two-tailed tests, and scare you with the heavens coming down on you if you were so reckless as to consider a one-tailed test.

Practitioners and academics alike were battling over which approach is best overall, or squabbling over the merits of particular applications - frequentist inference vs. decision-theoretic vs. Bayesian approaches. To make matters

even more confusing, there seemed to be noticeable schisms within each school of thought.

Most confusing of all, statistics as such turned out to be very context-dependent - it meant different things in different scientific and business fields. Practitioners in those fields had, over time, developed somewhat separate branches of statistics. Therefore, statistics would mean something different for you depending on whether you come from physics, medicine, social studies, econometrics, environmental studies, or industrial quality control.

It was simply a nightmare attempting to navigate this fractured jungle of jargon, conflicting stances, and math-heavy explanations. Yet, I persevered! And through painstaking reading, practice, implementing/coding methods, and countless simulation runs, I was able to garner a good enough understanding of the matter to begin writing methodological white papers and in-depth articles, to start delivering lectures and courses on statistics in A/B testing, and to become a developer of statistical tools.

From my current position, I see both the immense **value** of statistical methods applied to business risk management, estimation and prediction problems, and the immense **harm** done by improper applications or misguided understanding of those same methods. Thus, in-depth explanations of the practical application of statistical methods, as well as common errors and how to avoid them, are key elements of this work.

Furthermore, in 2019 the difficulties that I went through are about as severe as they were a few years before, despite the valiant efforts of some in the statistics and A/B testing communities. Addressing common mistakes, misconceptions, and misapplications of statistical methods is, therefore, a central part of this work.

My aim with this book is to **carve a clear path through the statistical jungle,** and thus save the reader weeks, months or even years of wandering around in circles, falling into gorges, and crossing rivers, metaphorically speaking! While the book does use the established jargon, each term is explained with painstaking detail and accuracy using the simplest language possible. Math and formulas are kept to a sanitary minimum in order to facilitate reading, while also satisfying the needs of technically-inclined readers, who will also find the detailed references supporting each chapter to be particularly useful. Since this is a book on statistical methods, and not on decision theory per se, the text also sticks to the frequentist error-statistical approach, and only briefly touches on current decision-theoretic and Bayesian methods.

WHO IS THE READER OF THIS BOOK?

This book aims to introduce the complex topic of statistical estimation and inference to readers with somewhere between little and no mathematical and statistical background. The text makes few assumptions, and builds each topic from the ground up, explaining the rationale behind each concept, and following it with a multitude of practical examples from the world of online A/B testing. It contains detailed explanations, so that one can understand the statistical methods deeply enough in order to correctly put them into practice, but steers clear of some of the difficult parts of set theory, calculus, etc., which are typical of many other books on statistics.

A background in conversion rate optimization, operating an online business or mobile app, design of user experiences, or similar, would be helpful for an easier reading, as the primary audience for this work is conversion rate optimization professionals who design, execute, and analyze A/B tests in an online environment. However, due to the similarities with other fields where controlled experimentation is possible and valuable, the book can be a useful guide to A/B testing in areas other than website and mobile application development. Product managers and growth experts should benefit from it regardless of the particular product or service they are focused on.

The overall framing of the presentation is always mindful of the topic of business objectives achieved through statistical methods. This will be useful for those readers who have some experience of statistics in other disciplines, and who are now looking to understand the use of statistical tools in facilitating decision-making through online A/B testing.

However, I must emphasize that the unique research presented in this book, in terms of ideas, models, and simulation results, will be valuable to all readers, regardless of background.