Statistical Methods in Online A/B Testing

Georgi Z. Georgiev

Statistical Methods in Online A/B Testing

This book is dedicated to explaining the tools of statistical inference and estimation through online controlled experiments; a.k.a. A/B tests. It views them in a risk-management context of balancing the risks and rewards of innovation. With the help of this text, user experience and conversion rate optimization practitioners will be able to harness the power of data-driven decision-making, and enable their business to innovate, while controlling the risk to which it is exposed.

An issue with much of the current statistical theory and practice in online A/B testing is that it is misguided and frequently misinterpreted. It is often applied without a good understanding of the role and limitations of statistical methods, instead blindly copying scientific applications without due consideration of many of the unique features of online business. This book approaches this problem by laying solid statistical foundations, and providing clear definitions and a multitude of practical examples, while constantly keeping an eye on the overarching business goal of A/B testing. By making constant use of the business context, and ample practical examples, this text presents A/B testing statistics in a uniquely useful way.

Georgi Z. Georgiev is the Managing Director of Web Focus LLC, and a veteran web marketer and web developer. His diverse 15-year experience includes owning, developing, and managing dozens of successful online projects, working as an SEO, Google AdWords, Google Analytics, and statistics consultant, as well as delivering training and lectures on multiple seminars and events, including in his capacity as a Google Regional Trainer. He is also developer of statistical tools at Analytics-Toolkit.com, as well as the author of numerous articles and white papers on the topic of statistics in online A/B testing. His most notable works have been “Efficient A/B Testing in Conversion Rate Optimization: The AGILE Statistical Method”, and the first extensive glossary of statistical terms in online A/B testing. His vast experience with online business and statistics positions him uniquely to deliver an accessible book on the topic of statistics applied to online A/B tests.

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.

This is the end of this free pdf preview of "Statistical Methods in Online A/B Testing". If you are looking for the eBook version of the book, it can be purchased on Kobo while the paperback is available on Amazon.

If you are looking for free works of mine, then please consider the vast amount I've already shared for free:

- Close to one hundred articles on the Analytics Toolkit blog, full of high-quality information.
- Several hundred entries of an A/B testing glossary which will get you up to speed with key statistical and experimentation concepts.
- Several whitepapers, available completely for free as well.

You are welcome to read all these free materials before you decide whether it is worth it for you to spend money on the book.

"Statistical methods in online A/B testing" has been the product of many months of effort which built on top of years of experience. If you are looking to obtain its value, you should do so by purchasing it on Amazon or Kobo - the two vendors I have authorized to distribute my work. This will help motivate me make more of my work available to the public one way or another.

Thank you.

Georgi Georgiev

Author of "Statistical Methods in Online A/B Testing"