Galen R. Shorack

Probability for Statisticians

With 23 Figures

Springer

Preface

There is a thin self-contained textbook within this larger presentation. To be sure that this is well understood, I will describe later how I have used this text in the classroom at the University of Washington in Seattle.

Let me first indicate what is different about this book. As implied by the title, there is a difference. Not all the difference is based on inclusion of statistical material. (To begin, Chapters 1–6, provide the mathematical foundation for the rest of the text. Then Chapters 7–8 hone some tools geared to probability theory, while Chapter 9 provides a brief introduction to elementary probability theory right before the main emphasis of the presentation begins.)

The classical weak law of large numbers (WLLN) and strong law of large numbers (SLLN) as presented in Sections 10.2–10.4 are particularly complete, and they also emphasize the important role played by the behavior of the maximal summand. Presentation of good inequalities is emphasized in the entire text, and this chapter is a good example. Also, there is an (optional) extension of the WLLN in Section 10.6 that focuses on the behavior of the sample variance, even in very general situations.

Both the classical central limit theorem (CLT) and its Lindeberg and Liapunov generalizations are presented in two different chapters. They are first presented in Chapter 11 via Stein's method (with a new twist), and they are again presented in Chapter 14 using the characteristic function (chf) methods introduced in Chapter 13. The CLT proofs given in Chapter 11 are highly efficient. Conditions for both the weak bootstrap and the strong bootstrap are developed in Chapter 11, as is a universal bootstrap CLT based on light trimming of the sample. The approach emphasizes a statistical perspective. Much of Section 11.1 and most of Sections 11.2–11.5 are quite unusual. I particularly like this chapter. Stein's method is also used in the treatment of U-statistics and Hoeffding's combinatorial CLT (which applies to sampling from a finite population) in the optional Chapter 17. Also, the chf proofs in Section 14.2 have a slightly unusual starting point, and the approach to gamma approximations in the CLT in Section 14.4 is new.

Both distribution functions (dfs $F(\cdot)$) and quantile functions (qfs $K(\cdot) \equiv F^{-1}(\cdot)$) are emphasized throughout (quantile functions are important to statisticians). In Chapter 7 much general information about both dfs and qfs and the Winsorized variance is developed. The text includes presentations showing how to exploit the inverse transformation $X \equiv K(\xi)$ with $\xi \cong \text{Uniform}(0, 1)$. In particular, Chapter 7 inequalities relating the qf and the Winsorized variance to some empirical process results of Chapter 12 are used in Chapter 16 to treat trimmed means and L-statistics, rank and permutation tests, sampling from finite populations, and bootstrapping. (Though I am very fond of Sections 7.6–7.11, their prominence is minimized in the subsequent parts of the text.)

At various points in the text choices can be voluntarily made that will offer the opportunity for a statistical example or foray. (Even if the instructor does not exercise a particular choice, a student can do so individually.) After the elementary introduction to probability theory in Section 9.1, many of the classical distributions of statististics are introduced in Section 9.2, while useful linear algebra and the multivariate normal distribution are the subjects of Section 9.3 and Section 9.4. Following the CLT via Stein's method in Section 11.1, extensions in Section 11.2–11.3, and application of these CLTs to the bootstrap in Sections 11.4–11.5, there is a large collection of statistical examples in Section 11.6. During presentation of the CLT via chfs in Chapter 14, statistical examples appear in Sections 14.1, 14.2, and 14.4. Statistical applications based on the empirical df appear in Sections 12.10 and 12.12. The highly statistical optional Chapters 16 and 17 were discussed briefly above. Also, the conditional probability Sections 8.5 and 8.6 emphasize statistics. Maximum likelihood ideas are presented in Section A.2 of Appendix A. Many useful statistical distributions contain parameters as an argument of the gamma function. For this reason, the gamma and digamma functions are first developed in Section A.1. Section A.3 develops cumulants, Fisher information, and other useful facts for a number of these distributions. Maximum likelihood proofs are in Section A.4.

It is my hope that even those well versed in probability theory will find some new things of interest.

I have learned much through my association with David Mason, and I would like to acknowledge that here. Especially (in the context of this text), Theorem 12.4.3 is a beautiful improvement on Theorem 12.4.2, in that it still has the potential for necessary and sufficient results. I really admire the work of Mason and his colleagues. It was while working with David that some of my present interests developed. In particular, a useful companion to Theorem 12.4.3 is knowledge of quantile functions. Sections 7.6–7.11 present what I have compiled and produced on that topic while working on various applications, partially with David.

Jon Wellner has taught from several versions of this text. In particular, he typed an earlier version and thus gave me a major critical boost. That head start is what turned my thoughts to writing a text for publication. Sections 8.6, 19.2, and the Hoffman–Jorgensen inequalities came from him. He has also formulated a number of exercises, suggested various improvements, offered good suggestions and references regarding predictable processes, and pointed out some difficulties. My thanks to Jon for all of these contributions. (Obviously, whatever problems may remain lie with me.)

My thanks go to John Kimmel for his interest in this text, and for his help and guidance through the various steps and decisions. Thanks also to Lesley Poliner, David Kramer, and the rest at Springer-Verlag. It was a very pleasant experience.

This is intended as a textbook, not as a research manuscript. Accordingly, the main body is lightly referenced. There is a section at the end that contains some discussion of the literature.

Contents

Chapter 7. Distribution and Quantile Functions

Chapter 8. Independence and Conditional Distributions

Chapter 9. Special Distributions

Chapter 10. WLLN, SLLN, LIL, and Series

Chapter 11. Convergence in Distribution

Chapter 12. Brownian Motion and Empirical Processes

Chapter 13. Characteristic Functions

Chapter 14. CLTs via Characteristic Functions

Chapter 15. Infinitely Divisible and Stable Distributions

Chapter 16. Asymptotics via Empirical Proceses

Chapter 17. Asymptotics via Stein's Approach

Chapter 18. Martingales

Chapter 19. Convergence in Law on Metric Spaces

Appendix A. Distribution Summaries