Preface

Roughly speaking, to solve an inverse problem is to recover an object (e.g., parameter or function) from noisy (typically indirect) observations. In most cases such recovery cannot be done exactly because the mathematical models that link data to the object are approximations, data are noisy, the number of observations is finite, and obtaining a solution may require further approximations for efficient numerical computations. The importance of assessing the reliability of solutions to inverse problems is evident given such potential sources of errors. This assessment step is part of what is now called uncertainty quantification (UQ). Uncertainty quantification for inverse problems and other problems in engineering requires familiarity with some basic methods from mathematics, probability, and statistics. But what I have observed during years of collaborations with scientists and applied mathematicians working on inverse problems is that they often do not feel as comfortable with their knowledge of probability or statistics as they do with their background in applied mathematics. The converse is also true: I have encountered statisticians interested in making contributions to inverse problems but who have not been exposed to the basic theory of inverse problems and the questions that arise in their applications. The objective of this book is therefore to serve as a bridge between the applied mathematics and statistics communities. I try to take advantage of the reader's mathematical background to provide a basic introduction to probability and statistics for UQ mainly in the context of inverse problems, a field with many important practical applications. In addition, the book provides a basic introduction to statistical regularization of inverse problems for those with a background in statistics. Since the reader is assumed to be comfortable with mathematical methods at the level of senior undergraduates and beginning graduate students in mathematics, engineering, and physical sciences, much ground can be covered: from undergraduate statistics and probability to probability distributions on infinite-dimensional spaces. For statisticians, the book uses classic linear regression and statistical inference to introduce the framework of ill-posed inverse problems and explain statistical questions that arise in their applications. A review of the mathematical analysis tools required for inverse problems is also included in the appendix. Since the statistical and probability methods covered have applications beyond inverse problems, the book may also be of interest to people working in data science or in other applications of UQ.

The selection of topics I cover has been strongly influenced by discussions I have had over the years with scientists, applied mathematicians, engineers, and students from a wide variety of fields. In particular, since advances in computational power have made the use of Bayesian methodology commonplace in many fields of application, I believe that the existence of different schools of inference to conduct UQ is a topic that deserves more attention. For example, I have encountered practitioners who were either not aware of the existence of non-Bayesian (e.g., frequentist) methods for inverse problems or could not tell (or care about) the difference. And among those who know about frequentist and Bayesian methods, there are many who are baffled by common heated discussions among statisticians regarding the merits or demerits of these two schools of inference. In practical applications we learn that Bayesian and frequentist methods (as well as likelihood methods) provide valuable tools for UQ and for statistical analysis in general. I therefore try to cover both frameworks and to explain their assumptions, corresponding interpretations, and their important complementary roles by means of examples. To keep the mathematics and probability theory accessible to a wide audience, I consider probability distributions mostly on finite-dimensional spaces but do provide some background and examples that serve as introduction to the infinite-dimensional case. However, even within the framework of finite-dimensional inverse problems, there is no single statistical methodology that will work in every application: UQ is highly problem dependent. I have chosen a particular framework that is widely used, has many practical applications, and provides basic tools for more complex problems.

We are all aware of how difficult it is to put to use new definitions and results as this requires techniques that are learned with experience. To help with this transition, each section includes examples with explicit calculations that introduce useful problem solving techniques relevant to the particular topic. Examples are also used to clarify theoretical concepts and to illustrate the type of applications for which the methods could be used. I include over 130 examples but choosing them has not been easy. I have tried to select simple illustrative examples that can be understood by a diverse audience. Although it may not be apparent, many of the examples are simplifications that capture the essence of more complex questions that arise in applications but which would require much background to explain fully. Some examples are in fact answers to questions I have received from students and collaborators over the years. Some sections also include more theoretical but important details to help warn the reader of subtle statistical/probabilistic issues that arise in applications of UQ and which could be easily overlooked.

The book is organized as follows. Chapter 1 provides an introduction to inverse problems and regularization. Chapters 2 and 4 cover probability and statistical methods whose applications to inverse problems are considered in Chapters 3, 4, and 5. Chapter 3 includes methods for data analysis, Chapter 4 focuses on Bayesian methods that are relevant to inverse problems, and Chapter 5 is dedicated to the data analysis of one particular set of experimental data. One of the goals of Chapter 5 is to illustrate the nuances that arise when we try to apply theory to the analysis of real data. The book includes two appendices: In order to make the book as self-contained as possible, and to establish the general terminology used throughout, Appendix A provides a summary of results from analysis that are used in different parts of the book. Given the importance of conditional probability for Bayesian inference, Appendix B provides a more careful discussion of conditional probability and conditional expectation, including the definition of regular conditional probability. Appendix B assumes some knowledge of measure theoretic probability but is not required for the understanding of the other chapters. It includes an introduction to an alternative approach to conditional probability based on disintegration which is not commonly taught. I believe this is a natural approach that may help some readers get a more intuitive understanding of conditional probability and expectation.

As explained above, the objective of this book is to provide a basic background in statistics and probability for UQ mainly in the framework of inverse problems. My hope is that this book can be used to complement other textbooks that focus on regularization, mathematical analysis or computational methods for inverse problems. Readers interested in learning more about regularization or the general theory of inverse problems may consult [93, 120, 150, 154], or [8, 126, 256] for more applied or computational introductions. The book [30] provides an edited collection of research papers and tutorials for UQ and large-scale inverse problems. For more material on Bayesian methods for finite-dimensional inverse problems see, for example, [49, 148, 246], and [241] for an introduction to Bayesian methods in the infinite-dimensional setting. The book should also help the reader learn the basic theory needed to study Markov chain Monte Carlo (MCMC) methods, which play a key role in Bayesian statistics but which I do not cover in this book. There are many good references dedicated to the theory or implementations of MCMC methods and Bayesian computation [42, 49, 113, 209, 234, 245]. The analysis of inverse problems also requires numerical optimization methods not discussed in this book. The reader may find introductions to optimization methods that are important for inverse problems in [8, 39, 192, 256]. Readers interested in learning more about general statistical methods, frequentist or Bayesian, may consult, for example, [22, 52, 55, 104, 163, 164, 224, 226]. The reader may also find [223] interesting as it provides a historical account of the role statistics has played in the twentieth century.

I would like to thank Fadil Santosa and the IMA for organizing and funding workshops where preliminary versions of this work have been used. I am also thankful to Roger Ghanem for giving me the opportunity to do the same at USC. I am very grateful to mentors, reviewers, colleagues, students, and friends who have either provided valuable feedback in the writing of this book, or who have played an important role in my understanding and appreciation of the subject. In particular, I am specially thankful to Oscar Aguilar, Vaughn Ball, Wolfgang Bangerth, Julianne and Tia Chung, Maarten de Hoop, Colin Fox, Mahadevan Ganesh, Eldad Haber, Alex Kalmikov, Paul Martin, Youssef Marzouk, Joyce McLaughlin, Bill Navidi, Aaron Porter, Juan Restrepo, John Scales, George Smoot, Alessio Spantini, and Philip Stark. Finally, I would like to thank Cheryl for her unfailing support and patience.

An apology regarding notation. Different areas of statistics, probability, mathematics, and physics have different notational conventions. For example, it is common in statistics to denote random variables with capital letters and their realizations with lower case (e.g., x is a value the random variable X takes). But in this book we need letters to denote sets, σ-algebras, random sets, scalars, vectors, matrices, random variables, random elements, functions, operators, measures, inner product spaces, normed spaces, etc. This makes it very difficult to follow any particular convention consistently. I hope the notation will be clear from the context.