**Cuốn 2**

Probability and statistics are fascinating subjects on the interface between mathematics and applied sciences that help us understand and solve practical problems.*(Xác suất thống kê là những môn học hay ho nằm giữa giao diện môn toán và khoa học ứng dụng mà nó giúp cho chúng ta hiểu được và giải được các bài toán trong thực tế)* We believe that you, by learning how **stochastic** *(ngẫu nhiên)* methods come about and why they work, will be able to understand the meaning of **statistical statements** *(phát biểu thống kê)* as well as judge *(đánh giá)* the quality of their content, when facing such problems on your own. Our **philosophy***(triết lý)* is one of how and why: instead of just presenting stochastic methods as **cookbook recipes** *(triết lý của chúng giải thích tại sao và như thế nào thay vì chỉ trình bày công thức như “mì ăn liền”),* we prefer to explain the principles behind them *(chúng tôi thích giải thích những nguyên tắc cơ bản ẩn đằng sau).* In this book you will find the basics of probability theory and statistics. In addition, there are several topics that go somewhat beyond the basics (ngoài xa còn có các chủ đề vượt xa chủ đề cơ bản) but that ought to be present in an introductory course: simulation (mô phỏng), the Poisson process, the law of **large numbers** (luật số lớn), and the central limit theorem. Computers have brought many changes in statistics. In particular, the bootstrap has earned its **place** (vị trí, chỗ đứng). It provides the possibility to derive confidence intervals and perform tests of hypotheses where traditional (normal approximation or large sample) methods are **inappropriate** *(không phù hợp)*. It is a modern useful tool one should learn about, we believe.

Examples and datasets in this book are mostly from real-life situations, at least that is what we looked for in **illustrations** (minh họa) of the material (dữ liệu trong cuốn sách này ở trong thực tế, ít nhất ở trong điều chúng tôi tìm kiếm để mình họa cho tài liệu này). Anybody who has inspected datasets with the purpose of using them as elementary examples knows that this is hard: *(bất kì ai bị dữ liệu ảnh hưởng với mục đích sử dụng chúng để ví dụ được biết thì khó)* on the one hand, you do not want to boldly state assumptions that are clearly not satisfied *(nói 1 cách khác, bạn không muốn những giả định trạng thái rõ rang mà không được đáp ứng)*; on the other hand, long explanations concerning side issues distract from the main points *(những lời giải thích dài dòng liên quan đến những vấn đề phụ làm sao nhãn từ những quan điểm chính)*. We hope that we found a **good middle way***(1 cách ở giữa)*. A first course in calculus is needed as a prerequisite for this book *(Giải tích 1 là môn học tiên quyết cho cuốn sách này).* In addition to high-school algebra, some infinite series are used (exponential, geometric). *(Bên cạnh đó môn đại số ở trung học, những chuỗi vô hạn hiện hữu trong)*.Integration and differentiation *(tích phân và vi phân)* are the most important skills, mainly concerning one variable *(chủ yếu liên quan đến hàm 1 biến)* (the exceptions, two dimensional integrals, are encountered *(bắt gặp)* in Chapters 9–11). Although the mathematics is kept to a minimum, we strived *(cố gắng)* to be mathematically correct throughout the book *(Chúng tôi cố gắng giảm toán những vẫn chính xác về mặt toán học)*. With respect to probability and statistics the book is self-contained (Đối với xstk là mọi thứ đều chứa đủ). The book is aimed at undergraduate engineering students, and students from more business-oriented studies *(Cuốn sách này nhắm tới sv học ngành kĩ thuật, và những sv hướng đến kinh tế)* (who may **gloss over** (lướt qua) some of the more mathematically oriented parts). At our own university we also use it for students in applied mathematics (where we put a little more emphasis on the math and add topics like combinatorics *(toán hỗ hợp)*, conditional expectations (giá trị kì vọng có điều kiện), and generating functions *(những hàm tạo)*). It is designed for a one-semester course (1 học kì): on average two hours in class per chapter, the first for a lecture (giảng bài lý thuyết), the second doing exercises. The material is also well-suited for self-study, as we know from experience.

We have divided attention about evenly between probability and statistics. The very first chapter is a sampler with differently flavored introductory examples, **ranging from** *(trải dài từ)* scientific success stories to a **controversial** *(phổ biến)* puzzle. Topics that **follow** *(theo sau)* are elementary probability theory *(lý thuyết xác suất cơ bản)*, simulation, joint distributions, the law of large numbers, the central limit theorem, statistical modeling (informal *(nôm na)*: why and how we can draw inference *(suy luận)* from data), data analysis, the bootstrap, estimation (đo lường), simple linear regression *(hồi quy tuyến tính cơ bản),* confidence intervals, and hypothesis testing. Instead of a few chapters with a long list of discrete (rời rạc) and continuous distributions, with an enumeration *(liệt kê)* of the important attributes of each, we introduce a few distributions when presenting the concepts and the others where they arise (more) naturally. A list of distributions and their characteristics *(đặc trưng)* is found in Appendix A