CHAPTER 1 – INTRODUCTION

TABLE OF CONTENTS

The Problem and Objectives

The material in this book derives from the simple observation that the available information is most often insufficient to provide a unique answer or solution for most interesting decisions or inferences we wish to make. In fact, insufficient information—including limited, incomplete, complex, noisy and uncertain information—is the norm for most problems across all disciplines.

The pervasiveness of insufficient information across the sciences has resulted in the development of discipline-specific approaches to dealing with it. These different approaches provide different insights into the problem. They also provide grist for an interdisciplinary approach that leverages the strengths of each. This is the core objective of this book. Here, I develop a unified constrained optimization framework – I call it *info-metrics* – for information processing, modeling and inference for problems across the scientific spectrum. The interdisciplinary aspect of this book provides new insights and synergies between distinct scientific fields. It helps create a common language for scientific inference.

Info-metrics combines the tools and principles of information theory, within a constrained optimization framework, to tackle the universal problem of insufficient information for inference, model and theory building. In broad terms, info-metrics is the discipline of scientific inference and efficient information processing. This encompasses inference from both quantitative and qualitative information, including nonexperimetal information, information and data from laboratory experiments, data from natural experiments, the information embedded in theory, and fuzzy or uncertain information

from varied sources or assumptions. The unified constrained optimization framework of info-metrics helps resolve the major challenge to scientists and decision makers of how to reason under conditions of incomplete information.

In this book I provide the mathematical and conceptual foundations for info-metrics and demonstrate how to use it to process information, solve problems, and construct models or theories across all scientific disciplines. I present a framework for inference and model or theory building that copes with limited, noisy, and incomplete information. While the level and type of uncertainty can differ among disciplines, the unified info-metrics approach efficiently handles inferential problems across disciplines using all available information. The info-metric framework is also suitable for constructing and validating new theories and models, using observed information that may be experimental or non-experimental. It also enables us to test hypotheses about competing theories or causal mechanisms. I will show that the info-metrics framework is logically consistent and satisfies all important requirements. I will compare the info-metrics approach with other approaches to inference and show that it is typically simpler, and more efficient to use and apply.

Info-metrics is at the intersection of information theory, statistical methods of inference, applied mathematics, computer science, econometrics, complexity theory, decision analysis, modeling, and the philosophy of science. In this book, I present foundational material emerging from these sciences as well as more detailed material on the meaning and value of information, approaches to data analyses, and the role of prior information. At the same time, this primer is not a treatise for the specialist; I provide a discussion of the necessary elementary concepts needed for understanding the methods of

info-metrics and their applications. As a result, this book provides even researchers with minimal quantitative skills with the necessary building blocks and framework to conduct sophisticated info-metric analyses.

With this book, I aim to provide a reference text that elucidates the mathematical and philosophical foundations of information theory and maximum entropy, generalizes it, and applies the resulting info-metrics framework to a host of scientific disciplines. The book is interdisciplinary and applications-oriented. It provides all the necessary tools and building blocks for using the info-metrics framework for solving problems, making decisions, and constructing models under incomplete information. The multi-disciplinary applications provide a hands-on experience for the reader. That experience can be enhanced via the exercises and problems at the end of each chapter.

Box 1.1. A Concise Historical Perspective

Outline of the Book

The plan of the book is as follows. The current chapter is an introductory one. It expresses the basic problem and describes the objectives and outline of the book. The next three chapters present the building blocks of info-metrics. Chapter 2 provides the rationale for using constrained optimization to do inference on the basis of limited information. This chapter invokes a specific decision function to achieve the kind of inference we desire. It also summarizes the axioms justifying this decision function. Despite the axiomatic discussion, this is a non-technical chapter. Readers familiar with constrained optimization and with the rational of using the entropy as the decision function may even skip this chapter.

The next two chapters present the mathematical framework underpinning the building blocks of Chapter 2. Chapter 3 explores the basic metrics of info-metrics; additional quantities will be defined in later chapters. Chapter 4 formulates the inferential problem as a maximum-entropy problem within the constrained optimization framework of Chapter 2, which is then formulated as an unconstrained optimization. Chapter 4 also develops the methods of validation to evaluate the inferred solutions.

The next two chapters provide a mix of detailed cross-disciplinary applications illustrating the maximum-entropy method in action. They demonstrate its formulation, its simplicity, and its generality in real world settings. Chapter 5 starts with a relatively simple set of problems. Chapter 6 presents more advanced problems and case studies.

Chapter 7 develops some of the basic properties of the info-metrics framework. It builds directly on Chapter 4 and concentrates on the properties of efficiency, optimality, and sufficiency. Chapter 7 fully quantifies the notion of 'best solution.'

Having formulated the basic building blocks the book moves on to the broader, more general info-metrics framework. Chapter 8 introduces the concept of prior information and shows how to incorporate such information into the framework. This chapter also takes up the critical question of how to construct this prior information, and it explores three different routes. The first approach is based on the grouping property – a property of the Boltzmann-Gibbs-Shannon entropy defined in Chapter 3 – which is less familiar to social and behavioral scientists. The second approach is based on the more obscure concept of transformation groups. Finally, the chapter considers empirical priors – a familiar but often misused concept for social scientists. Chapter 8 places special emphasis on the extension of these ideas to common problems in the social sciences.

Chapter 9 extends all previous results to accommodate all types of uncertainties, including model and parameters uncertainties. This chapter provides the complete info-metrics framework. All applications and specific problems can be modeled within the complete framework. It encompasses all inferential and model construction problems under insufficient information. Chapter 9 fully develops the complete interdisciplinary vision of this book and the complete info-metrics framework. The examples throughout the book complement that vision.

Combining the ideas of Chapter 9 with those of the earlier chapters takes us to model and theory building, causal inference, and the relationship between the two. The fundamental problem of model development and theory building is the subject of Chapter 10. The premise of this chapter is that the info-metrics framework can be viewed as a 'meta-theory' – a theory of how to construct theories and models given the imperfect information we have. That framework provides a rational perspective that helps us to identify the elements needed for building a reasonably sound model. That premise is demonstrated via multi-disciplinary examples, one of which is very detailed. The same building blocks are used to construct each one of these examples. This chapter also places emphasis on the idea that a model should be constructed on all of the information and structure we know or assume, even if part of that information is unobserved. In such cases a mechanism for connecting the observable information to the unobserved entities of interest must be provided. Examples of such a mechanism are discussed as well.

In Chapter 11 the emphasis shifts from model and theory building to causal inference via constraint satisfaction. The term *causal inference* here is taken to mean the causality inferred from the available information; we infer that A causes B from

information concerning the occurrences of both. The first part of this chapter concentrates on nonmonotonic and default logic, which were developed to deal with extremely high conditional probabilities. The second part deals with cause and effect in a probabilistic way given the information we have and the inferential framework we use. The chapter also provides detailed example that connects some of the more traditional ideas of causal inference to the info-metrics framework.

The next two chapters connect the info-metrics framework with more traditional statistical methods of inference. In particular, they show that the family of information-theoretic methods of estimation and inference are subsumed within the info-metrics framework. These chapters use duality theory to connect info-metrics with all other methods. Chapter 12 concentrates on discrete models. In that setting, specific maximum-likelihood approaches are special cases of the info-metrics framework. Chapter 13 concentrates on continuous models, such as linear and nonlinear regression analysis and system of equations analysis. It compares the info-metrics framework with the familiar least squares and other method of moments approaches for continuous models. Chapter 13 also shows that the info-metrics framework can accommodate possible misspecifications in empirical models. Misspecification issues are common across the social and behavioral sciences, where the researcher does not have sufficient information to determine the functional form of the structure to be inferred. Chapter 13 also demonstrates, via familiar examples, the tradeoffs between functional forms (the constraints in our framework) specification and the decision function used in the inference. To demonstrate this, the chapter shows that two different formulations yield the same inferred distribution even though one of the two is misspecified.

Chapter 14 provides four detailed, cross-disciplinary applications developed especially for this book. These applications represent diverse fields of investigation: the medical sciences, political science, and finance. The chapter illustrates the generality and simplicity of the info-metrics approach, while demonstrating some of the features discussed throughout the book. Each case study presents the required empirical background, the necessary analytics conditional on the input information, the inferred solution, and a brief summary of its implications.

Each chapter includes exercises and extended problems. Each chapter ends with a notes section, which summarizes the main references to that chapter as well as readings on related topics. The book is complemented by a website: http://info-metrics.org that provides supporting codes and data sets (or links to the data) for many of the examples presented in the book. It also provides extended analyses of some of the examples as well as additional examples.

A simple logical dependencies chart of the chapters is provided below. It shows the flow of the book. Though I recommend reading the chapters in their order, the diagram helps those that may be more informed or are just interested in a certain topic or problem. It also provides the necessary details for instructors and students.

[Insert Figure 1.1 (logical dependencies) here]

Legend for Figure 1.1:

Figure 1.1. Chapters' Logical Dependency Chart. The chart provides the logical flow of the book. Though there are many examples throughout the book, the three chapters devoted solely for examples shown above can be read in order (see arrows) or at any time after reading the relevant chapters.