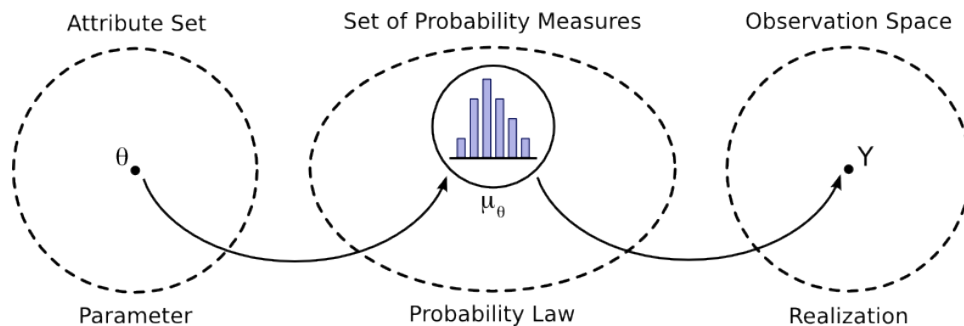


Chapter 1

Fundamentals of Statistical Signal Processing

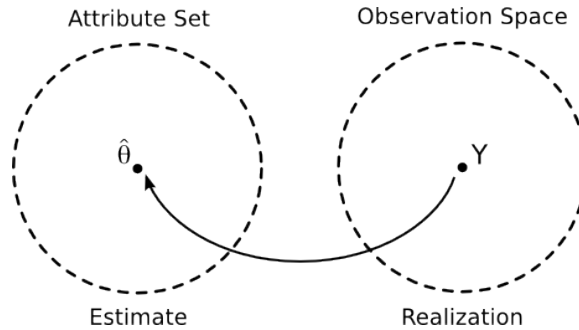
Statistical signal processing refers to the act of inferring generalizations from empirical data, with varying degrees of certainty. In this introduction, we provide a unified view of detection and estimation theory. The statistical problem we wish to explore can be described as follows. We are interested in an unknown *parameter* (or *phenomenon*) which we denote by θ . We have access to an observation Y that provides partial information about the value of θ . The connection between Y and θ is probabilistic in nature, rather than direct. In particular, different values of θ may produce the same observation. This abstract framework is illustrated below.



Three basic components are involved in this framework. First, there is the *attribute set* which is composed of all admissible values of θ . This set is

represented by U . The second component is a measurable space (\mathcal{F}, Ω) where Ω symbolizes a *universal sample space* and \mathcal{F} is the σ -algebra of all probability events. Together, \mathcal{F} and Ω provides a mathematical basis for the randomness present in the observations. Finally, the *observation space* Γ is the collection of all realizable observations. If the attribute set contains a unique element, then the inference problem is trivial and we need not observe Y to determine the value of θ . The more interesting case, of course, occurs when U contains several distinct elements. In this scenario, the connection between θ and Y becomes instrumental. We explore this relationship next.

To each element $\theta \in U$, there corresponds a probability law μ_θ acting on the measurable space (\mathcal{F}, Ω) . The true value of θ thus determines the distribution of random variable Y . While observing the realization of Y may provide information about its distribution, this empirical data is typically insufficient to resolve all the uncertainty surrounding the value of θ . Statistical signal processing broadly refers to the collection of techniques and methods employed to extract information about the phenomenon of interest from the available data.



We provide a simple example below to illustrate how this framework applies to engineering problems.

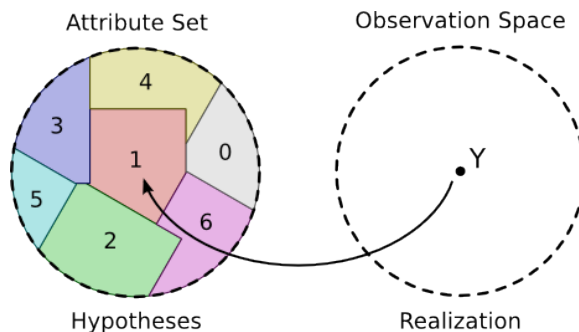
Example 1. *A digital communication system transmits a single bit over a noisy channel, using a zero or a one. The received information is corrupted by additive Gaussian noise, with distribution $\mathcal{N}(0, 1)$. We wish to decide which bit was sent from the received information. To initiate this process, we cast this problem in the abstract framework described above. In the present case,*

the attribute set is $U = \{0, 1\}$; the sample space is the real line; and the corresponding probability density functions are

$$f_0(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \quad f_1(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}}.$$

Note that in this first example the observation space and the sample space are identical.

The primary characteristics that distinguish statistical inference problems from one another are the amount of a priori knowledge available about the attribute set U , the goal underlying the inference procedure, and the performance criterion used to assess the effectiveness of the inference procedure. If the attribute set is partitioned into a finite number of subsets and the objective is to identify which subset θ belongs to, then the decision process is termed *detection* or *hypothesis testing*. In particular, a detector maps every element in the observation space Γ to one of the admissible hypotheses. A graphical interpretation of a detection process is illustrated below.



On the other hand, if U contains an infinite number of candidates and we are tasked with selecting the most appropriate one, then we are facing an *estimation* problem. In this setting, the goal is to select an estimate $\hat{\theta}$ based on observation Y as to optimize a given objective function. Thus, parameter estimation is a procedure that takes an argument in observation space Γ and returns an element in the attribute set U . The distinction between detection and estimation is at times rhetorical; we will not delve on this issue. Rather, our focus will be to develop a solid understanding of statistical signal

processing and to become well-versed in applying some of its techniques and algorithms.

A second important distinction among inference problems, as mentioned above, comes from the assumptions made on attribute set U . In *classical estimation*, U is taken to be a specific set containing the unknown parameter θ . This can be contrasted with the *Bayesian approach* where U is assumed to be a probability space with a specific distribution. In this latter case, the value of θ is itself the outcome of a random experiment. The distinction between these two approaches leads to disparate performance criteria, and hence different estimator. Such a distinction is also present in detection problems, where the attribute set is either a collection of object with a deterministic parameter θ , or a measurable space with a certain probability law.

A final and less straightforward distinction between inference problems stems from the nature of the empirical observations. The framework detailed above implicitly assumes that observations can be viewed as random variables. If on the other hand we have access to empirical data coming from a random process, then the inference problem becomes more challenging. In this scenario, inference problems are referred to as *signal detection* and *signal estimation*. When dealing with the progressive estimation of a stochastic process, the statistical inference procedure is often call *filtering*. We will address some of these problems once we acquire the mathematical sophistication necessary to tackle them.

1.1 Organization

These notes are organized as follows.

In Chapter 2, we study hypothesis testing. We begin with an introduction to the simplest such problem, binary detection. Both the classical and Bayesian forms of the problem are considered. The theory of detection is then extended to composite hypothesis testing, a framework where the attribute set contains more than two elements. We also discuss the problem of robust detection, sequential detection and quickest change detection. Finally, we look

at simple instances of M -ary hypothesis testing.

Chapter 3 is devoted to classical parameter estimation. We first introduce the celebrated Cramer-Rao lower bound, which provides a performance limit on unbiased estimators. We then present the concept of sufficient statistics. This is followed by a discussion of common estimators including the best linear unbiased estimator, the maximum likelihood estimator, and the least squares estimator. The expectation-maximization algorithm is also introduced as a numerically tractable estimation methodology.

Chapter 4 presents a study of Bayesian parameter estimation. The maximum a posteriori estimator is first introduced. This is followed by a discussion of the minimum mean square error and linear minimum mean square error estimators. Signal estimation is treated in Chapter 5, where the Kalman-Bucy filter and the Wiener-Kolmogorov filter are introduced.

1.2 Modeling Engineering Problems

These notes present a collection of important tools and methods in statistical signal processing. Becoming familiar with these tools should provide the reader with the ability to solve various inference problems, and to assess the performance of the corresponding statistical procedure. Still, an equally important skill to develop is the capacity to take a concrete engineering problem and to formulate it mathematically in a way that leads to a meaningful solution. In these notes, we will often use an abstract problem formulation as our starting point. However, the value of being able to take an engineering challenge and to cast it in a relevant yet solvable framework is not to be underestimated. Developing this skill is partly what sets apart the value of one's education and experience. Throughout, we will provide examples and case studies of meaningful problems together with suitable mathematical interpretations. Reading these examples carefully and understanding the process of selecting the proper tools to solve them should help develop good intuition. While going through these notes, one should also try to think about new situations where the statistical methods under consideration apply.