Chapter 1

Introduction

1.1. Definition

All models are wrong, but some are useful. Many statisticians know and appreciate G.E.P. Box's comment on statistical modeling (Box, 1979). Often the choice of the final inference model is a compromise of an accurate representation of the experimental conditions, a preference for parsimony and the need for a practicable implementation. However, these competing goals are not always honestly spelled out, and the resulting uncertainties are not fully described.

Over the last 20 years a powerful inference approach that allows us to mitigate some of these limitations has become increasingly popular. Bayesian nonparametric (BNP) inference allows us to acknowledge uncertainy about an assumed model while maintaining a practically feasible inference approach. We could take this feature as a pragmatic characterization of BNP as flexible prior probability models that generalize traditional models by allowing for positive prior probability for a very wide range of alternative models, while centering the prior around a parsimonious traditional model. A more formal definition of BNP is as probability models on infinite dimensional parameter spaces, such as functional spaces.

Example 1 (Density estimation) *Consider a simple random sample $y_i \sim F$ i.i.d., $i = 1, \ldots, n$, from some unknown distribution F. Bayesian inference requires that the model be completed with a prior for the unknown F in the sampling model. One could proceed by restricting F to a normal location family, $F = \mathsf{N}(\theta, 1)$. The model F is indexed by a finite dimensional parameter vector θ and the model is completed with a prior probability model for the finite dimensional θ. We are back to parametric Bayesian inference. Figure 1.1a shows the resulting inference conditional on an assumed random sample $\boldsymbol{y}$. Naturally, inference about the unknown F is restricted to the assumed normal location family and does not allow for multimodality or skewness. In contrast, a BNP model would proceed with a prior probability model $p(F)$ for the unknown distribution. Figure 1.1b contrasts the parametric inference with the flexible BNP inference under a Dirichlet process mixture prior.*

In Example 1 the infinite dimensional random quantity is an unknown distribution. Alternatively, the infinite dimensional quantity might be the unknown mean function $f(\cdot)$ in a regression problem, a response surface, a spectral density, or perhaps an autoregressive mean function in a nonparametric time series model. In the rest of these notes we will mostly focus on problems where the infinite dimensional quantity is an unknown probability measure $F(\cdot)$, as in example 1. The reason for this focus is simply tradition; most BNP models in the recent literature consider random probability measures.