

Preface

These notes arose out of a short course at UC Santa Cruz in summer 2010. Like the course, the notes provide an overview of some popular Bayesian nonparametric (BNP) probability models. The discussion follows a logical development of many commonly used nonparametric Bayesian models as generalizations of the Dirichlet process (DP) in different directions, including Pólya tree (PT) models, species sampling models (SSM), dependent DP (DDP) models and product partition models (PPM). The selection of topics is subjective, simply driven by what the authors are familiar with. As a result, some useful and elegant classes of models such as normalized random measures with random increments (NRMIs) are reviewed only briefly.

We focus on BNP models for random probability measures, keeping for example a discussion of Gaussian process priors to only a brief review in the introductory chapter. Also, we put the emphasis on developing models, rather than a discussion of BNP data analysis for important statistical inference problems. However, some data analysis is introduced by way of short examples and in an introductory chapter. Inference for BNP models often requires computation-intensive implementations. Keeping the focus on models, we decided against a discussion of computational algorithms at much length. The only exception are posterior simulation schemes for Dirichlet process (DP) and DP mixture models. Finally, we do not discuss asymptotic results. These are important and non-trivial. Excellent recent reviews appear in the monograph by Ghosh and Ramamoorthi (2003) and a review paper by Ghoshal (2010).

The notes start with an overview of discussed models, and a bit more, in Chapter 1. This overview serves the purpose of clarifying the relationships of the many models that are introduced later. We hope that the introduction will put the proliferation of BNP models in some perspective. But it also creates some repetition when some of the material that is already included in this initial overview is re-introduced later, reflecting also the nature of this manuscript as lecture notes. Figure 1.2 can serve as a one-figure overview of the rest of the notes. Chapter 2 motivates the upcoming long list of models by discussing some typical applications of BNP in data analysis. Then, Chapters 3 through 8 introduce some of the most popular BNP in more detail.

Finally, a word about notation. We generically use $p(\cdot)$ to indicate a probability model. The arguments clarify which model is meant. We use specific names for probability models only when the probability model itself is a random variable. For example $p(G)$ refers to a BNP model for the random probability measure G . We use boldface to distinguish vectors from scalars only when needed, for example $\mathbf{x} = (x_1, \dots, x_n)$, but usually do not use bold face when no confusion arises. Sometimes we use (x_i) to indicate a vector $(x_i, i = 1, \dots, n)$, when the range of the indices is clear from the context. Finally, we use notation like $\mathbf{N}(x \mid \mu, \sigma)$ to indicate a normal distributed random variable x with moments (μ, σ) . By a slight abuse of notation we use $\mathbf{N}(x \mid \mu, \sigma)$ also for the corresponding p.d.f.