Bayesian nonparametric analyses using Ferguson-Dirichlet process

Problems of statistical inference with an infinite dimensional parameter space, usually a space of probability distributions over a set, are of great importance both theoretically and practically. The Bayesian approach to such nonparametric problems requires that a probability distribution be placed over this space. The central class of distributions for use in these problems is the class of Dirichlet processes. (See Ferguson, Phadia, and Tiwari 1993)

First, we give Ferguson’s (1973) definition of the Dirichlet process and review briefly the developments in the basic theory of such process.

Let $A$ be a $\sigma$-field of subsets of a set $X$, and let $\alpha$ be a finite nonnull measure on $(X, A)$. A Dirichlet process $P$ with parameter $\alpha$, denote by $D(\alpha)$, if for all positive integers $k$ and every measurable partition $A_1, A_2, \ldots, A_k$ of $X$, the random vector $(P(A_1), \ldots, P(A_k))$ has a $k$-dimensional Dirichlet distribution with parameter $(\alpha(A_1), \ldots, \alpha(A_k))$. The basic result for this process is:

Theorem (Ferguson, 1973)
If $P$ is a Dirichlet process with parameter $\alpha$, and if, given $P$, $X_1, X_2, \ldots, X_n$ is a sample from $P$, then the posterior distribution of $P$ given $X_1, X_2, \ldots, X_n$ is a Dirichlet process with parameter $\alpha + \sum_{j=1}^{n} \delta(x_j)$, where $\delta(x)$ represents the distribution giving mass one to the point $x$.

A new construction simpler than that of Ferguson has been given in the following theorem.

Theorem (Sethuraman and Tiwari, 1982)
Let $Y_1, Y_2, \ldots$ be i.i.d. with beta distribution, $\text{Be}(M, 1)$, $M > 0$, let $Z_1, Z_2, \ldots$ be i.i.d. $F_0$, and let $\{Y_i\}$ and $\{Z_i\}$ be independent. Define $P_1 = (1 - Y_1)$, and $P_n = Y_1 \cdots (1 - Y_n)$ for $n > 1$. Then, $P = \sum P_j \delta(Z_j)$ is a Dirichlet process with parameter $\alpha = MF_0$.

Here, we shall use $M = \alpha(X)$ to represent the total mass of $\alpha$, and $F_0 = \alpha/M$ to be the prior guess at $P$.

Antoniak (1974) extends Ferguson’s result to cases where the random measure
is a mixing distribution for a parameter, which determines the distribution from
which observations are made. The conditional distribution of the random measure,
given the observations is no longer that of a simple Dirichlet process, but can be
described as being a mixture of Dirichlet process.

As a mean can be used to summarize a distribution, a random mean can be
used to summarize a random process. The distribution of the random mean of
Dirichlet process \( P \in D(MF_0) \), where \( F_0 \) is called the base measure (distribution)
and \( M(= \alpha(X)) \) is called precision parameter \( \int_X x dP(x) \) where \( X \) is the real line, has
been studied by several people, such as Hannum, Hollander, and Langberg (1981)
and Yamato (1984). Yamato (1984) shows that if \( F_0 \) is a Cauchy distribution, then
the random mean has the same Cauchy distribution. Jiang (1991) further considers
the case when \( X \) is a circle and shows that when \( M = 1 \) and \( F_0 \) is a uniform measure,
the random mean has a uniform distribution on the unit disk. Jiang, Dickey, and
Kuo (2004) extend the result to the case when \( X \) is an ellipse.

We shall further study random functional of the Dirichlet process when \( \alpha \) is
a different measure and/or \( X \) is a high dimensional space. The applications with
Dirichlet process or mixture of Dirichlet processes models have become increasingly
popular for modeling due to advances in simulation-based model fitting. Examples include empirical Bayes problems (Escobar, 1994); nonparametric regression
(Muller, Erkanli, and West 1996); density estimation (Escobar and West 1995; Gas-
parini 1996); hierarchical modeling (MacEachern 1994; West, Muller, and Escobar
1994; Bush and MacEachern 1996) and censored data settings (Doss 1995a; Kuo and
for Dirichlet process or mixture of Dirichlet processes models have been explored.
Some of them are given by Kuo (1986), MacEachern and Muller (1998), Neal (2000),
and Gelfand and Kottas (2002). Different computational approaches with possibly
more efficient estimation results shall also be explored. In addition, analyses of data
using the Dirichlet process models shall be investigated.
References


