A simple proof of the discreteness of Dirichlet processes

Nils Lid Hjort

Norwegian Computing Centre & University of Oslo

– August 1989, but now it's October 2021 –

Abstract. That Dirichlet processes are discrete with probability 1 is demonstrated once more.

Key Words: Dirichlet processes, discreteness

Let $\mathcal{X}$ be some sample space, and let P be a Dirichlet process with parameter $k\alpha$ on $\mathcal{X}$. Here k is a positive scalar and α is a probability measure on $\mathcal{X}$. Thus P is a random element of $\mathcal{M}$, the set of all probability measures on the sample space. There are various proofs in the literature that demonstrate that such a P is discrete with probability one; a possibly partial list includes Ferguson (1973, 1974), Blackwell (1973), Blackwell and MacQueen (1973), Doksum (1974), Kingman (1975), Berk and Savage (1979), Basu and Tiwari (1982), and Krasker and Pratt (1986). Here is yet another proof, from Hjort's 1976-thesis (p. 18). For a general overview of contemporary nonparametric Bayesian statistics, which includes the frequent use of Dirichlet processes in various forms, see Hjort (2003).

A useful lemma concerning the expected value of a function $f(P,X)$, where X is sampled from the random P, states that

$$\mathrm{E}\, f(P,X) = \int_{\mathcal{M}}\int_{\mathcal{X}} f(P,x)\, P(\mathrm{d}x)\, \mathcal{D}_{k\alpha}(\mathrm{d}P) = \int_{\mathcal{X}}\int_{\mathcal{M}} f(P,x)\, \mathcal{D}_{k\alpha+\delta_x}(\mathrm{d}P)\, \alpha(\mathrm{d}x).$$

Some measure theoretic details must be tended to here: there is some sigma-field $\mathcal{A}$ on $\mathcal{X}$; $\mathcal{M}$ is equipped, for example, by the Borel sets determined by set-wise convergence; and f must be measurable in (P,x). And $\mathcal{D}$, appropriately subscripted, is used to denote P's probability distribution.

The lemma is related in an obvious way to two well-known facts about the Dirichlet process, namely that X as above has unconditional distribution α, and that P, conditionally on an observed $X = x$, is Dirichlet with updated parameter $k\alpha + \delta_x$. There is a natural extension to functions $f(P, X_1, \ldots, X_n)$. The lemma was proved and used for various causes in Hjort (1976), and has later on been re-discovered on appropriate occasions; see Lo (1984) for but one example.

Introduce $A_P = \{x\colon P\{x\} > 0\}$, the set of atoms for a given P, and define

$$H_\gamma(P) = \mathrm{E}_P P\{X\}^\gamma = \int P\{x\}^\gamma\, \mathrm{d}P(x) = \sum_{x\in A_P} P\{x\}^{\gamma+1},$$

$$H_0(P) = \lim_{\gamma\to 0^+} H_\gamma(P) = \sum_{x\in A_P} P\{x\}.$$

A P is discrete if and only if $H_0(P) = 1$. Employ the lemma to get

$$\begin{aligned}\mathrm{E}\, H_\gamma(P) = \mathrm{E}\, P\{X\}^\gamma &= \int_{\mathcal{X}} \frac{\Gamma(k\alpha\{x\}+1+\gamma)}{\Gamma(k\alpha\{x\}+1)}\frac{\Gamma(k+1)}{\Gamma(k+1+\gamma)}\,\alpha(\mathrm{d}x)\\ &\ge \frac{\Gamma(1+\gamma)}{\Gamma(1)}\frac{\Gamma(k+1)}{\Gamma(k+1+\gamma)}.\end{aligned}$$

From this and $0 \le H_\gamma(P) \le H_0(P) \le 1$ follows $\mathrm{E}\, H_0(P) = 1$ and $H_0(P) = 1$ with $\mathcal{D}_{k\alpha}$-probability one. The single measure theoretic caveat here is that $(\mathcal{X}, \mathcal{A})$ must be such that $P\{x\}$ is simultaneously measurable. It suffices that $\mathcal{A}$ is the set of Borel sets from a metric which makes the sample space separable. Such conditions make H_γ measurable in P, and also entails the measurability of the set $\mathcal{M}_0$ of all discrete probability measures.

Arguments similar to those above show that each set A with positive α-measure must have positive P-atoms, with probability one. If in particular P is Dirichlet with parameter $k\alpha$ on the real line, with random distribution function F, then F has infinitely many jumps on each interval with positive α-measure.

Acknowledgments. This note has spent 45 years in Norwegian, and is based on pages 18–19 in my graduate thesis, written while an amanuensis vikar in Tromsø 1976. During that time I benefited from discussions with Berit Sandstad and Tore Schweder.

References

Basu, D. and Tiwari, R.C. (1982). A note on the Dirichlet process. In *Statistics and Probability: Essays in Honour of C.R. Rao*, eds. Kallianpur, Krishnaiah, Ghosh, 89–103.

Berk, R.H. and Savage, I.R. (1979). Dirichlet processes produce discrete measures: an elementary proof. *Contributions to Statistics: Jaroslav Hájek Memorial Volume*, Academia, North-Holland, 25–31.

Blackwell, D. (1973). Discreteness of Ferguson selections. *Ann. Statist.* **1**, 356–158.

Blackwell, D. and MacQueen, J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1**, 353–355.

Doksum, K. (1974). Tailfree and neutral random probability distributions and their posterior distributions. *Ann. Probab.* **2**, 183–201.

Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230.

Ferguson, T.S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2**, 615–629.

Hjort, N.L. (1976). *Applications of the Dirichlet process to some nonparametric estimation problems* (in Norwegian), graduate thesis, University of Tromsø. Abstracted in *Scand. J. Statist.* 1977.

Hjort, N.L. (2003). Topics in nonparametric Bayesian statistics (with discussion). In *Highly Structured Stochastic Systems* (eds. P.J. Green, N.L. Hjort and S. Richardson), Oxford University Press.

Kingman, J.F.C. (1975). Random discrete distributions (with discussion). *J. Royal Statist. Soc. Ser. B* **37**, 1–22.

Krasker and Pratt (1986). Contribution to the discussion of Diaconis and Freedman's 'On the consistency of Bayes estimates'. *Ann. Statist.* **14**, 42–45.

Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates: I. Density estimates. *Ann. Statist.* **12**, 351–357.

Sethuraman, J. and Tiwari, R. (1982). Convergence of Dirichlet measures and the interpretation of their parameter. In *Proceedings of the Third Purdue Symposium on Statistical Decision Theory and Related Topics* (eds. S.S. Gupta and J. Berger), 305–315. Academic Press, New York.