Dirichlet random probabilities and applications

Gérard Letac, Institut de Mathématiques de Toulouse

Abstract. If α is a fixed bounded measure on E, a Dirichlet random probability P governed by α is defined by the following property: if $E_0,\dots,E_n$ is a partition of E then the density of $(P(E_1),\dots,P(E_n))$ is proportional to

$$(1-x_1-\cdots-x_n)^{\alpha(E_0)-1}x_1^{\alpha(E_1)-1}\dots x_n^{\alpha(E_n)-1}.$$

This random measure is a very natural object in a number of contexts: non parametric Bayesian methods, solution of perpetuities equations, Markov-Krein theory, random paths. It is sometimes unappropriately called a Dirichlet process. The Italian school has made important contributions to its theory. The course will describe first the properties of P: even if α is continuous, P is almost surely purely atomic. It will study the existence and the distribution of the real random variable $\int_E f(w)P(dw)$, a particularly interesting object (For instance if $E=(0,1)$ and $\alpha(dx)=dx$ then the density of $X=\int_0^1 wP(dw)$ is proportional to $\sin(\pi x)x^{x-1}(1-x)^{-x}$). It will describe some of the applications mentioned above. The course necessitates no knowledge of stochastic integrals and martingales, but standard training in probability and measure theory is required, with a little dose of Lévy processes or infinitely divisible distributions. Up to this, it is fairly elementary.

A tentative schedule for eight classes of 2 academic hours is

1. The algebra of beta-gamma random variables. Dirichlet distributions, amalgamation property. The classical characterization of the gamma distributions. Classical objects of the exponential family of Dirichlet distributions.

2. The T_c transform of a probability on a tetrahedron: properties, examples and applications.

3. Definition and proof of the existence of the Dirichlet random probability (DRP). A DRP is a purely atomic distribution. Description of its random weights.

4. The Ewens distribution, the Chinese restaurant process. If P is DRP, the random variable $\int_E f(w)P(dw)$ exists if and only if

$$\int_E \log_+|f(w)|\alpha(dw)<\infty.$$

5. A short course on infinitely divisible distributions and on Lévy processes. The random measure associated to the Gamma process and its application to Dirichlet random probability.

6. Some particular cases of calculations of the distribution of $\int_{\mathbb{R}} wP(dw)$: when α is Cauchy, β and more generally uniform on a tetrahedron. The Markov-Krein moment problem. Applications to the non parametric Bayesian theory.

7. The Markov chains of the form $X_{n+1} = F_{n+1}(X_n)$ when $(F_n)_n \geq 1$ is an iid sequence of random maps from a set E into itself. Applications to perpetuities, examples when perpetuities are Dirichlet.

8. If Q is a probability on $\mathbb{R}^n$, $\theta > 0$ and if $Y \sim \beta(1, \theta)$, $W \sim Q$ and X are independent, then $X \sim (1 - Y)X + YW$ if and only if $X \sim \int xP(dx)$ where P is DRP governed by θQ. We prove this remarkable theorem of Diaconis and Freedman and extend it to $Y \sim \beta(k, \theta)$ with an other random measure called quasi Bernoulli of order k.

Some references

1. Cifarelli D.M. and Regazzini E. (1990) 'Distribution functions of means of a Dirichlet process ' *Ann. Statist.* **18**, 429-442.

2. Diaconis P. and Kemperman J. (1996) 'Some new tools for Dirichlet Priors' *Bayesian Statistics 5*, 97-106, Oxford University Press.

3. Diaconis P. and Freedman D. (1999) 'Iterated random functions' *Siam Review* **41**, pages 45-76.

4. James L.F., Lijoi A. and Prünster I. (2010) 'On the posterior distribution of classes of random means' *Bernoulli* **16**, 155-180.

5. Lijoi A. and Regazzini E. (2004) 'Means of a Dirichlet process and multiple hypergeometric functions' *Ann. Probab.* **32**, 1469-1495.