KdV institute	Nonparametric Bayesian Statistics, assignment set 1
UvA	September 25, 2014

- **Deadline: October 9, 2014.**
- **Send a pdf file with your answers and a text file with your R (or matlab) script to hvzanten@uva.nl.**
- **Your name and student number should be on the answer sheet!**

1. Let Θ be a Polish space, let $\mathcal{B}$ be its Borel σ-algebra, and let Π be a probability measure on $(\Theta, \mathcal{B})$. Prove that there exists a smallest closed set $F \subset \Theta$ such that $\Pi(F) = 1$.

2. Let P be a Dirichlet process on $\mathbb{R}$ with base measure α. Use the stick-breaking representation to prove that P has full support if and only if α has full support.

3. (a) In R (or Matlab, but R is preferred), write a simple script that can generate plots of the distribution function of a Dirichlet process on $\mathbb{R}$ with a given base measure α on a given grid. (Hint: use the gamma representation.)
   (b) Using the script, generate 3 pictures, each showing 10 realizations of (the distribution function of) the Dirichlet process, with base measures Φ, 10Φ and 100Φ, respectively, where Φ is the standard normal distribution.
   (c) What is the difference between the three pictures? Explain this from the theoretical properties of the Dirichlet process.

4. Let $P \sim DP(\alpha)$, with α a finite base measure on $\mathbb{R}$. Given P, let $X_1, \ldots, X_n$ be i.i.d., real-valued random variables with distribution P. Let ψ be a bounded, measurable function.
   (a) Compute the posterior mean and variance of $\int \psi \, dP$. (Hint: first consider $\psi = 1_A$.)
   (b) Prove that if the data are in actual fact sampled from the true distribution P_0, then as $n \to \infty$, the posterior distribution of $\int \psi \, dP$ tends to the Dirac measure concentrated at $\int \psi \, dP_0$ in an appropriate sense.