Dirichlet Distribution and Its Applications

Xi Tan (tan19@purdue.edu)

June 18, 2016

1 Gamma Function and Beta Function

The *gamma function* is defined for all complex numbers except the non-positive integers. For complex numbers with a positive real part, it is defined via an improper integral that converges:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \tag{1}$$

For all positive numbers z,

$$\Gamma(z) = (z-1)\Gamma(z-1) \tag{2}$$

and in particular, if n is a positive integer:

$$\Gamma(n) = (n-1)! \tag{3}$$

Note, the gamma function shifts the normal definition of factorial by 1.

The *beta function* is defined by

$$\mathrm{B}(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt \tag{4}$$

for $Re(x), Re(y) > 0$.

It can also be written as

$$\mathrm{B}(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{5}$$

2 Introduction

The Dirichlet distribution of order $K \geq 2$ with parameters $\alpha_1, \ldots, \alpha_K > 0$ has a probability density function with respect to Lebesgue measure on the Euclidean space $\mathbb{R}^{K-1}$ given by

$$f(x_1, \ldots, x_{K-1}; \alpha_1, \ldots, \alpha_K) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\alpha_i - 1} \tag{6}$$

I'm not quite sure about this "Lebesque measure on the Euclidean space" thing.

for all $x_1, \ldots, x_K > 0$ and $x_1 + \cdots + x_K = 1$. The density is zero outside this open $(K-1)$-dimensional simplex.

"Open" here means none of the x_i's can be 1, actually $x_i \in (0, 1)$.

The normalizing constant is the multinomial beta function, which can be expressed in terms of the gamma function:

$$B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{K} \alpha_i)}, \text{ where } \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K) \tag{7}$$