Probability Distributions: Continuous

Introduction to Data Science Algorithms

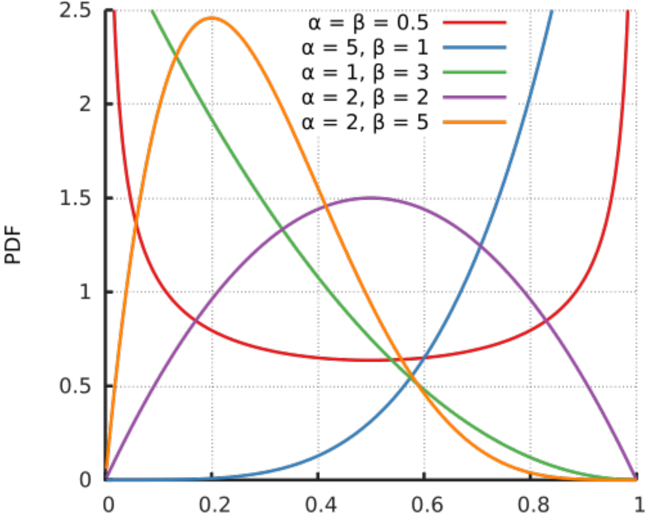
Jordan Boyd-Graber and Michael Paul

SEPTEMBER 14, 2016

Beta distribution

- The Beta distribution is over real values on the interval $[0,1]$.
- Useful distribution for modeling percentages and proportions
  - Batting averages in baseball
  - Percentage of people with a disease in a country
- The density is proportional to: $x^{\alpha-1}(1-x)^{\beta-1}$
- Related to the Bernoulli distribution: $x^{\theta}(1-x)^{1-\theta}$

Beta distribution



Beta distribution

- The PDF for Beta is:

$$f(x) = \underbrace{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}}_{\substack{\text{Inverse Beta function,}\\ \text{related to the binomial coefficient}}} x^{\alpha-1}(1-x)^{\beta-1}$$

- Γ is the gamma function, $\Gamma(x) = (x-1)!$
  - Just like the factorial function, but works for real values in addition to integers
- Mean: $\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$
- The parameters α and β must be > 0.

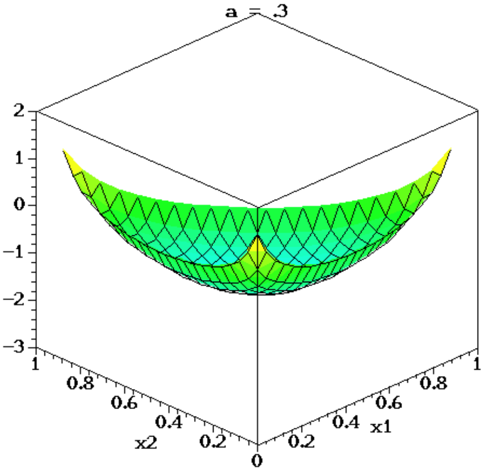
Dirichlet distribution

- The Dirichlet distribution is a generalization of the Beta distribution for multiple random variables
- The Dirichlet distribution is over *vectors* whose values are all in the interval $[0, 1]$ and the sum of values in the vector is 1.
  - In other words, the vectors in the sample space of the Dirichlet have the same properties as probability distributions.
  - The Dirichlet distribution can be thought of as a “distribution over distributions”.

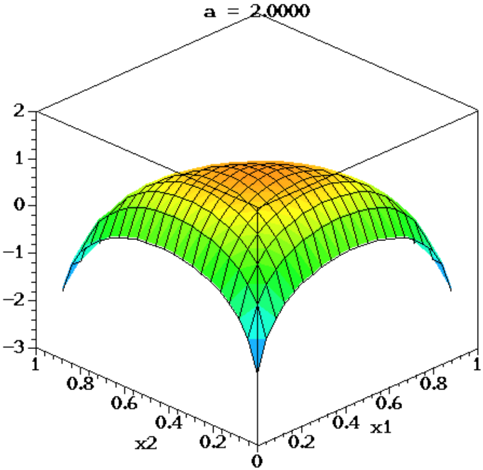
- The PDF for a K-dimensional Dirichlet distribution has a vector of parameters denoted α, given by:

$$f(\mathbf{x}) = \frac{\prod_{k=1}^{K} \Gamma(\alpha_k)}{\Gamma\left(\sum_{k=1}^{K} \alpha_i\right)} \prod_{k=1}^{K} x_k^{\alpha_k - 1}$$

Dirichlet distribution



Dirichlet distribution



Dirichlet distribution

- The Dirichlet PDF looks similar to the multinomial distribution.
    - The Dirichlet density is proportional to: $\prod_k x_k^{\alpha_k - 1}$
    - The multinomial mass is proportional to: $\prod_k x_k^{\theta_k}$

- Remember this analogy:
    - **Beta : binomial :: Dirichlet : multinomial**