Course Notes **Spring 2019**
Statistical Pattern Recognition (COMP136)
Rishit Sheth
Department of Computer Science, Tufts University

Some Probability Distributions

Beta distribution. The density for $\text{Beta}(\theta|a,b)$ where $\theta \in [0,1]$ is given by

$$p(\theta|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1},$$

where $\Gamma(x) = \int_0^\infty \tau^{x-1}e^{-\tau}\mathrm{d}\tau$ is referred to as the Gamma function. It is easy to verify that $\Gamma(1) = 1$. Additionally, one can show

$$\Gamma(x+1) = x\Gamma(x).$$

For an integer x, this implies that $\Gamma(x+1) = x!$.

Dirichlet distribution. The *multivariate* generalization of the Beta distribution to K categories is the Dirchlet distribution $\text{Dir}(\boldsymbol{\mu}|\boldsymbol{\alpha})$ where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)$. The density function is given by

$$p(\boldsymbol{\mu}|\boldsymbol{\alpha}) = \frac{\Gamma(\sum_{j=1}^K \alpha_j)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_K)}\prod_{j=1}^K \mu_j^{\alpha_j-1},$$

where

$$0 \leq \mu_j \leq 1, \qquad \sum_{j=1}^K \mu_j = 1.$$

These conditions ensure that the probability mass function represented by $\boldsymbol{\mu}$ is valid and are analogous to the condition on θ being constrained to $[0,1]$ in $\text{Beta}(\theta|a,b)$.

Univariate Gaussian distribution. The Gaussian (or Normal) density $\mathcal{N}(x|\mu,\sigma^2)$ where $x \in \mathbb{R}$ is given by

$$p(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$