Dirichlet Process

Aaron Schein

March 2, 2014

Table 1: Notational mapping

Description	Wallach	Murphy	Neal
Observation i	d_i	$\mathbf{x}_i$	y_i
Cluster index of i	z_i	z_i	c_i
Observation distribution (OD)	Multinomial(ϕ_{z_i})	$F(\theta_{z_i})$	$F(\phi_{c_i})$
Params of (OD)	ϕ_{z_i}	θ_{z_i}	ϕ_{c_i}
Params of OD prior	β, n	λ	
OD prior distribution	$P(\phi_k \vert \beta, n) = \text{Dir}(\phi_k \vert \beta, n)$	$\theta_k \sim H(\lambda)$	$\theta_c \sim G_0$
Mixing proportions (MP)	θ	π	p
Responsibility of cluster k	$P(z_i = k \mid \theta) = \theta_k$	$P(z_i = k \mid \pi) = \pi_k$	$P(c_i = k \mid p) = p_k$
Params of MP prior	α, m	α	α
MP prior distribution	$P(\theta \mid \alpha, m) = \text{Dir}(\theta; \alpha, m)$	$P(\pi \mid \alpha) = \text{Dir}(\pi \mid (\alpha/K)\mathbf{1}_k)$	$p_1, \cdots, p_K \sim \text{Dir}(\alpha/K, \cdots, \alpha/K)$