CS 228 Problem Set 5

Hugh Zhang

March 15, 2017

Problem 1

Note that from the hint, we can see that the Dirichlet distribution is a distributed beta, so all the lecture theorems apply.

$$P(X[M+1] = x^i \mid D)$$
$$= \sum_{\theta} P(X[M+1] = x^i \mid \theta) * P(\theta \mid D)$$
$$= \sum_{\theta} \theta_i * P(\theta \mid D)$$
$$= E[\theta_i \mid D] = \frac{\alpha_i + M[i]}{M + \alpha}$$

As proven in lecture. TODO GO TO OFFICE HOURS

1.2

$$P(X[M+1] = x^i, X[M+2] = x^j \mid D)$$
$$= P(X[M+1] = x^i \mid D) * P(X[M+2] = x^i \mid X[M+1] = x^j, D)$$

1.3

$$P(X[M+1] = x^i, X[M+2] = x^j \mid D)$$
$$- P(X[M+1] = x^i \mid D) * P(X[M+2] = x^i \mid D)$$
$$= \frac{\alpha_i + M[i]}{M + \alpha} * \frac{\alpha_j + M[j] + 1(i = j)}{M + \alpha + 1} - \frac{\alpha_i + M[i]}{M + \alpha} * \frac{\alpha_j + M[j]}{M + \alpha}$$
$$= \frac{\alpha_i + M[i]}{M + \alpha} \left(\frac{\alpha_j + M[j] + 1(i = j)}{M + \alpha + 1} - \frac{\alpha_j + M[j]}{M + \alpha} \right)$$
$$= \frac{\alpha_i + M[i]}{M + \alpha} \left(\frac{(M + \alpha)\alpha_j + (M + \alpha) * M[j] + (M + \alpha) * 1(i = j)}{(M + \alpha + 1) * (M + \alpha)} - \frac{(M + \alpha + 1)\alpha_j + (M + \alpha + 1) * M[j]}{(M + \alpha)(M + \alpha + 1)} \right)$$
$$= \frac{\alpha_i + M[i]}{M + \alpha} \left(\frac{-\alpha_j - M[j] + (M + \alpha) * 1(i = j)}{(M + \alpha + 1) * (M + \alpha)} \right)$$
$$=$$

The difference here is if you update your Dirichlet counts after you sample the M+1th sample when you do exact inference and don't if you do approximate. Thus, if your data set is large, then not updating doesn't matter, but if it is small it might make a large difference.

The numerator can at most asymptotically approach M^2, while the denominator is assymptotically M^3, thus the error asymptotically decreases linearly as M increases.