Morgan Ciliv

SPR Programming Project 1

Part 1:

As the training set size increases, the perplexities of the different methods converge to the same point. This phenomenon occurs because, the only variables that change with size are mk and N. Therefore, all of them will converge to the MLE, mk/N in the long run.

A shortcoming for the MLE for the unigram model is that it doesn't take into account alpha. The other estimates do take into account alpha, the model.

I think the full training set I think MLE will be the least sensitive. MAP will be sensitive, and the predictive distribution will be the most sensitive. I think this because of how each of the corresponding equations relate to alpha.