

Linguistics 696f
Hammond
Spring '03

Exercise #2

This is due by the beginning of class on Mar. 24.

- (1) Construct three bigram models *manually* on the basis of the first training text below:
 - a. an unsmoothed unigram model
 - a. an unsmoothed bigram model
 - c. a model interpolating these two models ($\lambda_1 = .7, \lambda_2 = .3$)
- (2) *Manually* calculate the cross-entropy of the second test text using the second and third models above.

Training text: "she sells sea shells by the sea shore"

Test text: "sea shells she sells by the she shore"