(a)

for word BY, the probality is 0.004180

for word BE, the probality is 0.003370

for word BUT, the probality is 0.002994

for word BEEN, the probality is 0.001658

for word BECAUSE, the probality is 0.000902

for word BILLION, the probality is 0.000822

for word B., the probality is 0.000684

for word BEFORE, the probality is 0.000666

for word BUSINESS,the probality is 0.000541

for word BUSH, the probality is 0.000516

for word BANK, the probality is 0.000464

for word BETWEEN, the probality is 0.000458

for word BEING, the probality is 0.000453

for word BACK, the probality is 0.000448

for word BASED, the probality is 0.000424

for word BOTH, the probality is 0.000400

for word BIG, the probality is 0.000340

for word BOARD, the probality is 0.000334

for word BEGAN, the probality is 0.000276

for word BILL, the probality is 0.000267

for word BLACK, the probality is 0.000239

for word BONDS, the probality is 0.000213

for word BEST, the probality is 0.000212

for word BETTER, the probality is 0.000212

for word BUY, the probality is 0.000209

for word BUDGET, the probality is 0.000205

for word BANKS, the probality is 0.000201

(b)

The word HUNDRED has probality 0.209061

The word <UNK> has probality 0.124304

The word .POINT has probality 0.099952

The word OF has probality 0.073947

The word THOUSAND has probality 0.068654

The word MILLION has probality 0.031832

The word ,COMMA has probality 0.031622

The word -HYPHEN has probality 0.030479

The word HALF has probality 0.029139

The word .PERIOD has probality 0.024376

(c)

Lu = -64.509

Lb = -40.918

The Lb is higher than Lu

(d)

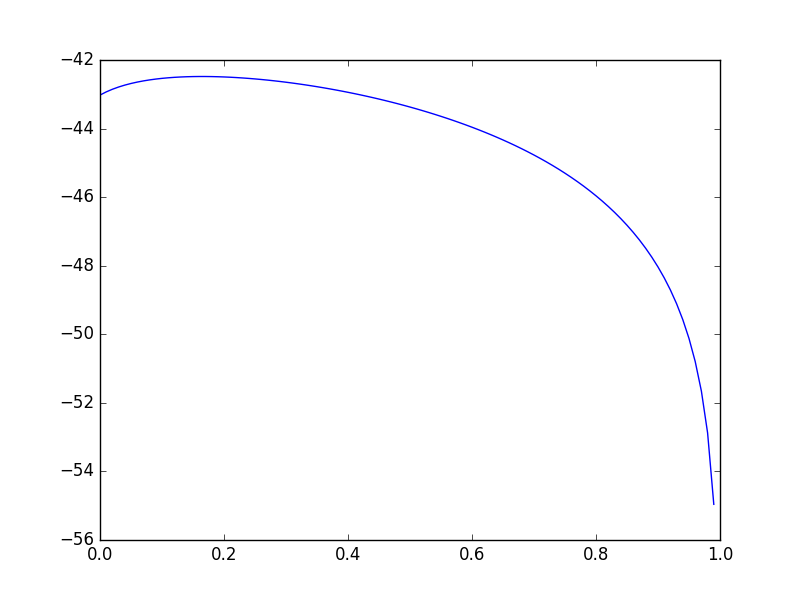
Lu = -44.2364299857

Lb = undefined,

no FOURTEEN followed by OFFICIALS

So the missing adjacent pair will cause the undefined problem in Lb

(e)



The highest should around 0.2