# CIS 3362 - 9/3/21

(1) IC - relevance finding keyword length

(2) MIC, relevance in reduce search for keyword

(3) QUIZ (Wed)

(2.5) tools for cryptanalysis you can use.

---

IC - probability of choose 2 of same item from a multi-set.

$$= \frac{\sum_{i=1}^{k} f_i(f_i-1)}{n(n-1)} \qquad n = \sum_{i=1}^{k} f_i .$$

IC (rnd english) ≈ .0378

IC ( english ) ≈ ~~.076~~ .0676

Ciphertext : $c_0, c_1, c_2, \ldots, c_{n-1}$

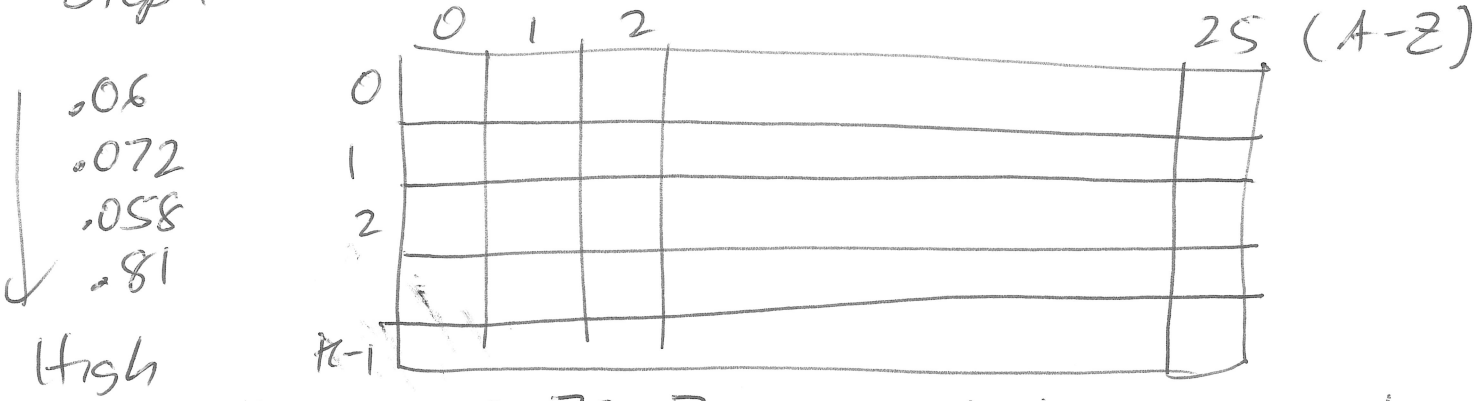
Guess k = keyword length — create k bins of letters

bin 0: $c_0, c_k, c_{2k}, c_{3k}, \ldots$

bin 1: $c_1, c_{k+1}, c_{2k+1}, \ldots$

bin i: $c_i, c_{k+i}, c_{2k+i}, \ldots$

Step 1: Create k bins



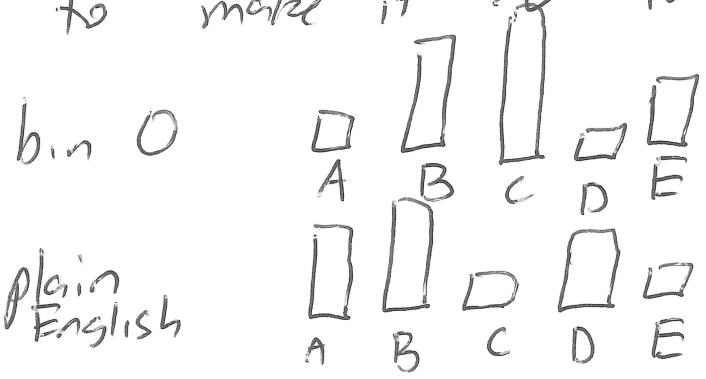

bin[i][j] = # of times character j appears in bin i.

Keep on trying different k's until the avg IC's is ~ .0676 (highest)

---

Step 2: What do I do once I have # of bins?

If we know the right k, then for each bin, I know some shift ought to make it ~ to English letter freq.




Given 2 freq arrays how do I test they are similar?

Mutal index of coincidence: Given two multisets of objects what's probability is you pull at random 1 item from each, that they are the same!

| | a | b | c | d | |
|---|---|---|---|---|---|
| Set 1 | 5 | 10 | 20 | 10 | 45 |
| Set 2 | 4 | 11 | 5 | 30 | 50 |

$$\frac{5}{45}\times\frac{4}{50}+\frac{10}{45}\times\frac{11}{50}+\frac{20}{45}\times\frac{5}{50}+\frac{10\times 30}{45\times 50}$$

$$=\frac{20+110+100+300}{45\times 50}=\frac{530}{45\times 50}=\frac{53}{45\times 5}$$

$$=\boxed{\frac{53}{225}}$$

Set 1 $f_1, f_2, \dots f_k$ $\quad n=\sum_{i=1}^{k} f_i$

Set 2 $g_1, g_2, \dots g_k$ $\quad m=\sum_{i=1}^{k} g_i$

$$MIC(Set1, Set2)=\frac{\sum_{i=1}^{k} f_i g_i}{nm}$$

| Example | Shift 0 | Shift 1 |
|---|---|---|
| bin 0 | 5, 20, 25, 5, 10 | 20, 25, 5, 10, 5 |
| P.E. | 20, 30, 5, 15, 5 | 20, 30, 5, 15, 5 |

try Shift 0 $= 5\times 20+20\times 30+25\times 5+5\times 15+10\times 5$
$= 950$

try shift 1 $= 20\times 20+25\times 30+5\times 5+10\times 15+5\times 5$
$400+750+25+150+25$
$=1350$

# Method #2

for each of the k bins try all 26 shifts compared w/ English Freq. find the shift that maximizes MIC for each bin and that should be your keyword.

# Method #1

① Brute force all 26 shifts on bin 0.

~~then k~~

② For bins $i =$ #1 to (k-1) do the followings:

Calc MIC (bin 0, bin $i$ Shifted by each value 0 to 25.

bin 1 shifted back 3 = bin 0
bin 2 shifted back 5 = bin 0
bin 3 shifted back 14 = bin 0

if bin 0's shift was A (0) key

A D F O
B E G P
C F H Q
D G I R
E H J S
etc.

# QUIZ

(1) No calc
(2) Ref Sheet – We will give
(3) Short Answer