2. Suppose we have a computer program which generates words from the alphabet A,B,C,D according to the following procedure:
   Pick the first letter according to the single frequency table given below then constructs each additional letter using the table of conditional biletter frequencies given below.
   a) Calculate the probability that the program produces the word "DACB"
   b) Determine the 2 letter word that has the highest probability.

Single letter table

A	10
B	9
C	12
D	9
	40

Biletter table

	A	B	C	D
A	0	3	1	0
B	2	1	4	0
C	0	2	0	3
D	2	1	0	1

$P(\text{word} = DACB) = P(\text{first letter} = D) \cdot P(\text{second letter} = A \mid \text{first letter} = D) \cdot P(\text{third letter} = C \mid \text{second letter} = A) \cdot P(\text{fourth letter} = B \mid \text{third} = C)$

$= \frac{9}{40} \cdot \frac{2}{4} \cdot \frac{1}{4} \cdot \frac{2}{5} = \frac{9}{800}$

AA —	BA $9/40 \cdot 3/7$	CA —	DA $9/40 \cdot 2/4$
AB $1/4 \cdot 3/4$	BB $9/40 \cdot 1/7$	CB $12/40 \cdot 3/5$	DB $9/40 \cdot 1/4$
~~AC $1/4 \cdot 1/4$~~	BC $9/40 \cdot 4/7$	CC —	DC —
AD —	BD —	CD $3/10 \cdot 3/5$	DD $9/40 \cdot 1/4$

$9/48 = 3 \cdot 3/3/16$ $<=?$ $9/50$

AB is largest probability

The index of coincidence is defined as

$$I_c = \frac{\text{number of pairs of equal letters in ciphertext}}{\text{the total number of pairs of letters}}$$

That is if we set

- N_α = the number of occurrences of the letter α in the cyphertext
- $$D_c = \sum_{\alpha=A}^{Z} \binom{N_\alpha}{2} = \sum_{\alpha=A}^{Z} \frac{N_\alpha(N_\alpha - 1)}{2}$$

  D_c represents the number of pairs of equal letters in the cyphertext.
- then $I_c = \frac{D_c}{\binom{N}{2}}$
- where N = the number of letters in the cyphertext