



INDEX OF COINCIDENCE

Index of coincidence (IC): is the probability that two letters selected from the text are identical, we can compute the index of coincidence (IC) from apply this formula:

$$IC = \frac{\sum f(f - 1)}{n(n - 1)}$$

Where f is the frequency of the letter in the text

n the length of the letter.

The IC is differs from one language to another. We can use the IC to discover if the message were enciphered using Monoalphabetic system or Polyalphabetic system.

To compute the key length (K), we apply this formula:

$$K = \frac{0.0256 * n}{(0.065 - IC) + n (IC - 0.0385)}$$

**Example 1**

If the A = 2; B = 16; C = 6; D = 3; E = 19,

What is the index of coincidence (IC).

Answer

$$IC = \frac{\sum f(f - 1)}{N(N - 1)}$$

$$f(f - 1)$$

$$A: 2(2-1) = 2$$

$$B: 16(16-1) = 240$$

$$C: 6(6-1) = 30$$

$$D: 3(3-1) = 6$$

$$E: 19(19-1) = 342$$

$$\sum f(f - 1) = 2 + 240 + 30 + 6 + 342 = 620$$

$$N = 2 + 16 + 6 + 3 + 19 = 46$$

$$IC = \frac{\sum f(f - 1)}{N(N - 1)} = \frac{620}{46(46 - 1)} = \frac{620}{46 * 45} = \frac{620}{2070} = 0.2995$$

**Example 2**

A sample of ordinary English contains the following distribution of letters,
Estimate the key word length.

Letter	Count	Letter	Count	Letter	Count
A	41	J	4	S	64
B	36	K	18	T	182
C	36	L	56	U	59
D	103	M	27	V	13
E	188	N	119	W	55
F	37	O	132	X	3
G	34	P	28	Y	23
H	102	Q	1	Z	0
I	123	R	95		

Solution

Apply the $f(f - 1)$ for every character as explain below

A	41(41-1)	= 1640	J	4 (4-1)	= 12	S	64 (64 -1)	= 4032
B	36 (36-1)	= 1260	K	18 (18 - 1)	= 306	T	182 (182 -1)	= 32942
C	36 (36-1)	= 1260	L	56 (56 -1)	= 3080	U	59 (59 -1)	= 3422
D	103 (103-1)	= 10506	M	27 (27 -1)	= 702	V	13 (13-1)	= 156
E	188 (188-1)	= 35156	N	119 (119 - 1)	= 14042	W	55 (55-1)	= 2970
F	37 (37-1)	= 1332	O	132 (132 - 1)	= 17292	X	3(3-1)	= 6
G	34 (34-1)	= 1122	P	28 (28 -1)	= 756	Y	23(23-1)	= 506
H	102 (102 -1)	= 10302	Q	1(1-1)	= 0	Z	0 (0-1)	= 0
I	123 (123 -1)	= 15006	R	95 (95 -1)	= 8930			



$$N = 1679$$

$$IC = \frac{\sum f(f-1)}{n(n-1)}$$

$$IC = \frac{1640 + 1260 + 1260 + 10506 + 35156 + 1332 + 1122 + 10302 + 15006 + 12 + 306 + \dots}{1679(1679-1)}$$

$$IC = \frac{184838}{1679 * 1678} = 0.0656$$

$$IC = \frac{184838}{2817362} = 0.0656$$

$$K = \frac{0.0256 * n}{(0.065 - IC) + n(IC - 0.0385)}$$

$$K = \frac{0.0256 * 1679}{(0.065 - 0.0656) + 1679(0.0656 - 0.0385)}$$

$$K = \frac{42.9824}{-0.0006 + 45.5009} = \frac{42.9824}{45.5003} = 0.94466$$



Example 3

A polyalphabetic ciphertext has the following letters counts.

Letter	Count	Letter	Count	Letter	Count
A	60	J	67	S	29
B	50	K	23	T	66
C	42	L	45	U	87
D	64	M	44	V	63
E	51	N	28	W	19
F	63	O	83	X	43
G	19	P	44	Y	39
H	48	Q	69	Z	67
I	56	R	13		

A	60(60-1)	= 3540	J	67 (67-1)	= 4422	S	29 (29 -1)	= 812
B	50 (50-1)	= 2450	K	23 (23 - 1)	= 506	T	66 (66 -1)	= 4290
C	42 (42-1)	= 1722	L	45 (45 -1)	= 1980	U	87 (87 -1)	= 7482
D	64 (64-1)	= 4032	M	44 (44 -1)	= 1892	V	63 (63-1)	= 3906
E	51 (51-1)	= 2550	N	28 (28 - 1)	= 756	W	19 (19-1)	= 342
F	63 (63-1)	= 3906	O	83 (83 - 1)	= 6806	X	43(43-1)	= 1806
G	19 (19-1)	= 342	P	44 (44 -1)	=1892	Y	39(39-1)	= 1782
H	48 (48 -1)	= 2256	Q	69(69-1)	= 4692	Z	67 (67-1)	= 4422
I	56 (56 -1)	= 3080	R	13 (13 -1)	= 153			

$N = 1282$



$$IC = \frac{\sum f(f-1)}{n(n-1)}$$

$$IC = \frac{3540 + 2450 + 1722 + 4032 + \dots}{1282(1282-1)}$$

$$IC = \frac{35761}{1282 * 1281} = \frac{35761}{821121} = 0.04355$$

$$K = \frac{0.0256 * n}{(0.065 - IC) + n(IC - 0.0385)}$$

$$K = \frac{0.0256 * 1282}{(0.065 - 0.04355) + 1282(0.04355 - 0.0385)}$$

$$K =$$

**H.W 1**

What is the index of coincidence for a collation of the letters consisting of the following? And find the key.

Letter	Count	Letter	Count	Letter	Count
A	65	J	67	S	29
B	0	K	20	T	66
C	42	L	40	U	87
D	64	M	44	V	63
E	51	N	30	W	19
F	60	O	80	X	40
G	19	P	41	Y	39
H	48	Q	69	Z	67
I	56	R	13		