

(F. Pratt, "Secret and Urgent", Blue Ribbon Books, 1942)

OCURRENCIA DE PARES DE LETRAS EN EL IDIOMA INGLES

Frecuencias aproximadas para 1,000 palabras:

	E	T	A	O	N	R	I	S	H	D	L	F	C	M	U	G	Y	P	W	B	V	K	X	J	Q	Z
E	14	14	31	4	61	88	7	2	2	46	18	9	17	18	3	7	7	9	4		10		9			
T	46	9	29	41		29	45	12	168		5	1	2	1	9		8		4							
A		68		1	92	40	16	42		21	33	3	18	9	6	11	9	8	3	8	10	7				
O	3	13	2	12	71	45	4	8	1	6	15	49	5	30	40		3	23	8	4	5	2		1		
N	31	37	13	11	5		16	24		61	4	3	19		1	27	6				2					
R	91	20	26	31	8	5	36	18		10	4		5	9	7	2	11	2	2	1	4	3				
I	14	38	16	31	86	16		41		13	17	3	24	18	1	9		8		3	13	3	2			1
S	35	63	5	19	1		21	14	14		1		6	2	14	1	3	5	2			3			1	
H	132	7	33	15	1	3	43			1	2				5		3		1							
D	32		8	4		5	14	6		1	2			2	8	1	1									
L	39	7	21	13		2	22	5		11	19	2		1	5		17		1		1					
F	9	2	7	17		12	11				2	9			3											
C	26	12	20	27		6	9		22		9		5		8		3					10				
M	27		27	14		2	13	3						5	4		3	9	6							
U	4	24	6		22	19	2	23		2	10	2	9	3		9		8		1						
G	12	1	5	6	1	7	5	3	12		2				2	4	1									
Y	4		3	4			1	4		1																
P	13	6	14	17		18	4	6	3		13			3	2			5								
W	29		22	10	3	2	15		16																	
B	23		10	10		5	2	2		9					7		7									
V	29		7	2			9										1									
K	8		2	1			5	3																		
X	3	3					1						1					3								
J	3		4	2											2											
Q															5											
Z	1																									
Σ	628	324	311	292	351	304	301	216	238	182	156	81	111	101	132	71	83	80	31	17	45	28	11	1	1	1

OCURRENCIA DE LETRAS EN EL IDIOMA INGLES

Frecuencia de ocurrencia en 1,000 palabras:

E	T	A	O	N	R	I	S	H	D	L	F	C	M	U	G	Y	P	W	B	V	K	X	J	Q	Z	Σ
591	473	368	360	320	308	286	275	237	171	153	132	124	114	111	90	89	89	68	65	41	19	7	6	5	3	4505

GRUPOS DE LETRAS (26 letras ordenadas por frecuencia relativa)

- I. E
- II. T
- III. A, O, N, R, I, S
- IV. H
- V. D, L, F, C, M, U
- VI. G, Y, P, W, B
- VII. V, K, X, J, Q, Z

PROBLEMA:

Suponga que las letras son emitidas por una fuente de Markov de primer orden. Calcule $H(S)$ utilizando la hoja excel que se adjunta. Compare su resultado con el valor dado en el texto (3.32 bits/símbolo).