- Entropy of random letters with probability of each letter chosen $1/26$

$$\sum_{\alpha=A}^{Z} 1/26 \log_2(26) \approx 4.7$$

- Entropy of letters in English chosen independently using single letter probabilities p_α

$$\sum_{\alpha=A}^{Z} p_\alpha \log_2(1/p_\alpha) \approx 4.16$$

- Entropy of English using biletter statistics ≈ 3.2
- Experimental entropy of English with at least 25 letters ≈ 1.2

Let F represent the of number of bits of information per English letter of text.

General rules	$n < 8$	$F = 4.16$
	$8 < n \leq 15$	$F = 3.2$
	$15 < n \leq 25$	$F = 2$
	$25 < n$	$F = 1.2$

For a cyphertext only attack to estimate the unicity distance (set $H(K|C) = 0$) we use the equation

$$H(C) = H(K) + H(M)$$

hence

$$n4.7 = H(K) + nF$$

If we solve for n we conclude

$$\boxed{n = \frac{H(K)}{4.7 - F}}$$

For a known plaintext attack we have the following theorem.

Theorem

$$H(K|C,M) = H(K) - H(C|M)$$

Proof: Applying the identity $H(X,Y) = H(X) + H(Y|X)$ we have

$$H(K,C,M) = H(C,M) + H(K|C,M) = H(M) + H(C|M) + H(K|C,M)$$

on the other hand

$$H(K,C,M) = H(K,M) = H(K) + H(M)$$

since K and M are independent. Hence

$$H(M) + H(C|M) + H(K|C,M) = H(K) + H(M)$$

and solving for $H(K|C,M)$ yields the identity.