# Vigenere Cipher, Kasiski and Friedman Attacks

## **Kasiski attack or Kasiski test (based on the textbook: *Making Breaking Codes, At Introduction to Cryptology*, by Paul Garret)**

* The goal is to determine the key length
* The idea is that if two trigrams in the plaintext occur at a distance apart which is a multiple of key length, then they will encrypt to the same trigram in ciphertext.
* The method: we look fro trigrams which occur more than once in the ciphertext, and speculate that their distances apart may be multiples of the keylength.
* The implementation:
	+ For each trigram in the ciphertext that occurs more than once, we compute the GCD of the collection of all distances apart of its occurrences.
	+ If GCD > 1 than we list the trigram and the associated GCD.
	+ We would guess that the keyword length is a divisor of at least one of these great common divisors.
* Problems:
	+ if the ciphertext is too small, then the chances are small that the same trigram will occur twice or more at a distance apart which happens to be a multiple of the key
	+ if the ciphertext is too large, the chances are greater that identical trigrams appear in siphertext for other reasons
1. **Friedman Attack or Friedman Test**

**(based on the textbook: *Making Breaking Codes, At Introduction to Cryptology*, Paul Garret and Invitation to Cryptology, by Thomas H. Barr)**

* The goal is to find a key length
* This attack is affective in determining the key length for any periodic substitution cipher
* The Index of Coincidence:

Given two streams of characters

Y = (y0, y1, y2, ….yN)

Z = (z0, z1, z2, …..zN)

The index of coincidence is



* If two streams are identical, I(Y, Z) = 1,
* If one (or both) of two streams are completely random, the index of coincidence would be expected to be 1/26 = 0.038 – if all the letters are random, then the probability is 1/26 that the two characters at the *i* - th place will match. So the index would be expected to be:

I(Y, Z) = (1/26 +1/26 +….1/26)/N = N\*(1/26)/N = 1/26 = 0.0385

* Different formula for index of coincidence I(Y, Z):
	+ Let n0, n1, …n25 be the respective frequencies of letters A, B…..Z in the ciphertext
	+ n = n0+n1+…n25 be the total number of letters in the text.



* The key length could be found by using the following formula:

