What does any of this entropy stuff have to do with cryptography?

Example

Language with 4-letter alphabet: \{a, b, c, d\}

\[ P(a) = .7, \quad P(b) = P(c) = P(d) = .1 \]

Individual letters are independent

Let \( M^n \) denote the random variable consisting of a typical \( n \)-gram from this language.

\[ M^n = (M_1, M_2, \ldots, M_n) \]

Similarly \( C^n \) is a random \( n \)-gram of ciphertext.
\[ H(M^1) = -[0.7 \cdot \log_2(0.7) + 3 \cdot 0.1 \cdot \log_2(0.1)] = 0.36 + 0.997 = 1.36. \]

Since plaintext characters are independent, \[ H(M^n) = n \cdot H(M^1) = 1.36n. \]
Consider a simple substitution cipher.

There are $4! = 24$ possible ciphers. Let $K$ (the key) denote a randomly chosen cipher.

$$H(K) = \log_2 24 = 4.6$$ since all keys are equally probable.

Since each character of ciphertext is equally likely,

$$H(C^1) = \log_2 4 = 2.$$ 

Are distinct ciphertext characters independent? . . . No!
\[ P(C^2 = xx) = 0.13 \quad \text{for } x \in \{a, b, c, d\} \]
\[ P(C^2 = xy) = 0.04 \quad \text{with } x \neq y. \]

Therefore

\[ H(C^2) = \]
\[ - [4 \cdot 0.13 \cdot \log_2(0.13) + 12 \cdot 0.04 \cdot \log_2(0.04)] = \]
\[ 1.53 + 2.23 = 3.76. \]
What does $H(K \mid C^n)$ represent?

Suppose we get ciphertext: ddadbd

Surely $E_K(a) = d$. We have narrowed down the set of possible keys from 24 to 6.

$H(K \mid C^n)$ is called the key equivocation.
**Theorem:** For any cryptosystem, 

\[ H(K | C^n) = H(K) + H(M^n) - H(C^n). \]

In our example:

\[
\begin{align*}
H(K) &= \log_2 24 = 4.6 \\
H(C^1) &= 2 \\
H(C^2) &= 3.76 \\
H(M^1) &= 1.36 \\
H(M^2) &= 2 \cdot 1.36 = 2.72
\end{align*}
\]

\[
\begin{align*}
H(K | C^1) &= 4.6 + 1.36 - 2.0 = 3.96 \\
H(K | C^2) &= 4.6 + 2.72 - 3.76 = 3.56
\end{align*}
\]