1 A Priori and A Posteriori Probabilities

Model Scenario

Consider

- a finite set $M_0 \subseteq M$ of possible plaintexts—for example all plaintexts of length r or of length $\leq r$,
- a finite set K of keys,
- a cipher $F = (f_k)_{k \in K}$ with $f_k \colon M \longrightarrow \Sigma^*$.

The restriction to a finite set M_0 allows us to handle probabilities in the naive way. It is no real restriction since plaintexts of lengths $> 10^{100}$ are extremely unlikely in this universe that has at most 10^{80} elementary particles.

Motivating Example

For English plaintexts of length 5 we potentially know exact a priori probabilities, say from a lot of countings. A small excerpt from the list is

Plaintext	Probability
hello	$p > 0$
fruit	$q > 0$
xykph	0
...	...

Now assume we see a monoalphabetically encrypted English text `XTJJA`. Without knowing the key—that is in a situation where all keys have the same probability—and without further context information we nevertheless assign to the single plaintexts different "a posteriori probabilities":

Plaintext	Probability
hello	$p_1 >> p$
fruit	0
xykph	0
...	...

Thus knowledge of the ciphertext alone (and knowledge of the encryption method) changed our information on the plaintext.

A "BAYESian" approach gives a general model of this observation.

Model

The probability of plaintexts is given as a function

$$P \colon M_0 \longrightarrow [0,1] \quad \text{where} \quad P(a) > 0 \quad \text{for all } a \in M_0$$
$$\text{and} \quad \sum_{a \in M_0} P(a) = 1.$$

(This is the a priori probability of plaintexts.)

The probability of keys is likewise given as a function

$$P\colon K \longrightarrow [0,1] \quad \text{such that} \quad \sum_{k\in K} P(k) = 1.$$

(By abuse of notation denoted by the same letter P.) In general we assume a uniform distribution $P(k) = 1/\#K$ for all $k \in K$.

The probability of ciphertexts derives from the probabilities of plaintexts and keys, implicitly assumed as independently chosen:

$$P\colon \Sigma^* \longrightarrow [0,1], \quad P(c) := \sum_{a\in M_0} \sum_{k\in K_{ac}} P(a)\cdot P(k),$$

where $K_{ac} := \{k \in K \mid f_k(a) = c\}$ is the set of all keys that transform a to c.

Remark 1 Only finitely many $c \in \Sigma^*$ have $P(c) \neq 0$. These form the set

$$C_0 := \{c \in \Sigma^* \mid P(c) > 0\}$$

of "possible ciphertexts".

Remark 2 We have

$$\begin{aligned}
\sum_{c\in\Sigma^*} P(c) &= \sum_{c\in\Sigma^*} \sum_{a\in M_0} \sum_{k\in K_{ac}} P(a)\cdot P(k) \\
&= \sum_{a\in M_0} \sum_{k\in K} P(a)\cdot P(k) \\
&= \sum_{a\in M_0} P(a)\cdot \sum_{k\in K} P(k) \\
&= 1.
\end{aligned}$$

The conditional probability for a ciphertext to stem from a given plaintext $a \in M_0$ is modeled by the function

$$P(\bullet|a)\colon \Sigma^* \longrightarrow [0,1], \quad P(c|a) := \sum_{k\in K_{ac}} P(k).$$

Remark 3 $\sum_{c\in\Sigma^*} P(c|a) = \sum_{k\in K} P(k) = 1.$

Remark 4 $P(c) = \sum_{a\in M_0} P(a)\cdot P(c|a).$

A Posteriori Probabilities of Plaintexts

The cryptanalyst is interested in the converse, the conditional probability $P(a|c)$ of a plaintext $a \in M_0$ if the ciphertext $c \in \Sigma^*$ is given.

First we describe the probability of the simultaneous occurrence of a and c as

$$P\colon M_0 \times \Sigma^* \longrightarrow [0,1], \quad P(a,c) := P(a) \cdot P(c|a).$$

Remark 5 Then

$$\sum_{a \in M_0} P(a,c) = \sum_{a \in M_0} P(a) \cdot P(c|a) = P(c).$$

The conditional probability of a plaintext is given by a function $P(\bullet|c)$ with $P(a,c) = P(c) \cdot P(a|c)$ by the BAYESian formula

$$P(a|c) := \begin{cases} \frac{P(a) \cdot P(c|a)}{P(c)} & \text{if } P(c) \neq 0, \\ 0 & \text{if } P(c) = 0. \end{cases}$$

Remark 6 $\sum_{c \in \Sigma^*} P(c) \cdot P(a|c) = \sum_{c \in \Sigma^*} P(a) \cdot P(c|a) = P(a)$ by Remark 3.