

## **Empirical Distribution**

## 1 Why

## 2 Definition

The **empirical distribution** of a data set is the distribution which assign probabilities to elements of the base set in accordance to their proportions in the data set.

## 2.1 Notation

Let  $(a^1, \ldots, a^n)$  be a data set in A. Let R be the set of real numbers. Let  $q: A \to R$  be defined by

$$q(a) = \frac{1}{n} |\{k \in \{1, \dots, n\} | a^k = a\}|$$

Then q is the empirical distribution of  $(a^1, \ldots, a^n)$ .