Empirical Law

1 Why

Suppose we have collected data.

2 Definition

Let A be a non-empty set. Let n be a natural number. A *data set* of size n for A is a function from $\{1, \ldots, n\}$ into A. It may be that $a_i = a_j$ for some $i \neq j$.

To each data set we associate an *empirical law* which is a probability measure P on the measurable space $(A, 2^A)$ that assigns to each set $B \subset A$ the number

$$P(B) = \frac{|\{i \in [n] \mid a_i \in B\}|}{n},$$