$$\diamondsuit$$

Empirical Measure

1 Why

There is a natural probability measure on a measurable space to associate with a dataset from the base set of that space.

2 Definition

The *empirical measure* for a dataset in some measurable space is the measure which associates to each event the proportion of the records which are elements of that event.

2.1 Notation

Let $(a^1, \ldots, a^n)$ be a dataset in a measurable space $(A, \mathcal{A})$. Let $P : A^* \to [0,1]$ be the probability measure that assigns to each set $B \subset A$ the number

$$P(B) = \frac{1}{n} \left| \left\{ k \in \{1, \ldots, n\} \mid a^k \in B \right\} \right|.$$

Subset Algebras

Cardinality

2

Injective Functions