Why

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

Definition

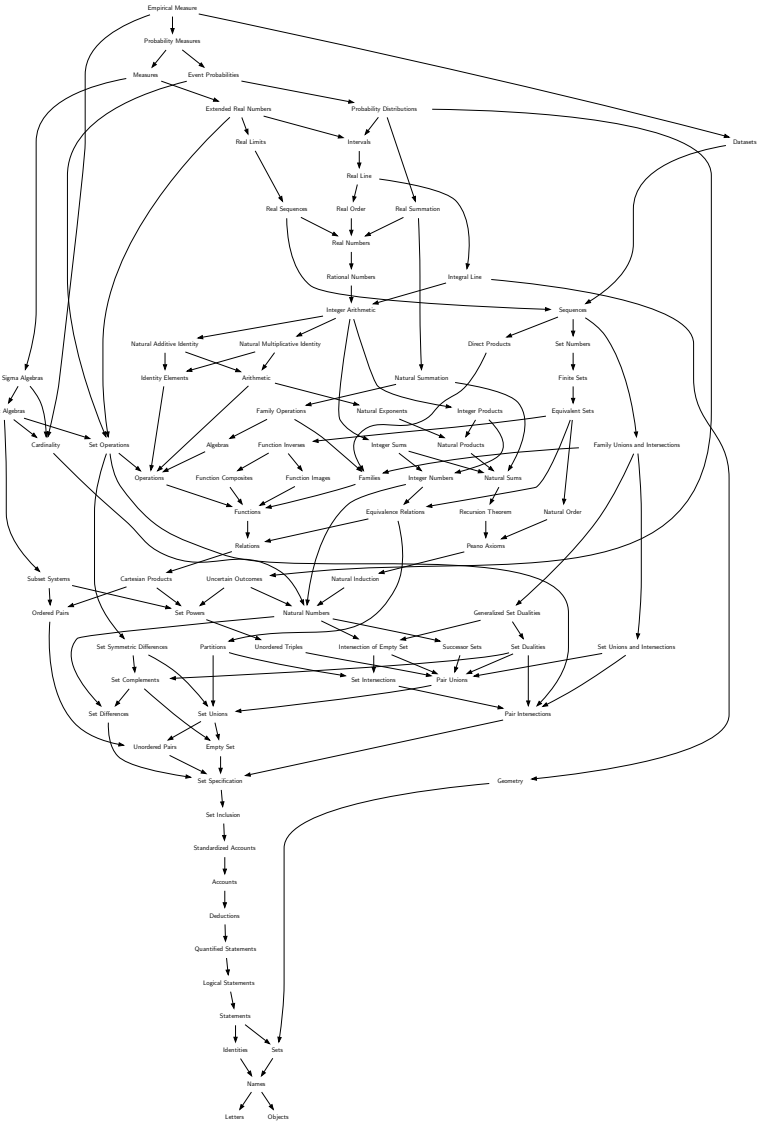
The *empirical measure* for a dataset in some measurable space is the measure which associates to each event the proportion of the records which are elements of that event.

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

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

Then P is the empirical measure.


Empirical Measure
Probability Measures
Measures
Event Probabilities
Extended Real Numbers
Probability Distributions
Datasets
Real Limits
Intervals
Real Line
Real Sequences
Real Order
Real Summation
Real Numbers
Rational Numbers
Integral Line
Integer Arithmetic
Sequences
Natural Additive Identity
Natural Multiplicative Identity
Direct Products
Set Numbers
Sigma Algebras
Identity Elements
Arithmetic
Natural Summation
Finite Sets
Subset Algebras
Family Operations
Natural Exponents
Integer Products
Equivalent Sets
Cardinality
Set Operations
Algebras
Function Inverses
Integer Sums
Natural Products
Family Unions and Intersections
Operations
Function Composites
Function Images
Families
Integer Numbers
Natural Sums
Functions
Equivalence Relations
Recursion Theorem
Natural Order
Relations
Peano Axioms
Subset Systems
Cartesian Products
Uncertain Outcomes
Natural Induction
Ordered Pairs
Set Powers
Natural Numbers
Generalized Set Dualities
Set Symmetric Differences
Partitions
Unordered Triples
Intersection of Empty Set
Successor Sets
Set Dualities
Set Unions and Intersections
Set Complements
Set Intersections
Pair Unions
Set Differences
Set Unions
Pair Intersections
Unordered Pairs
Empty Set
Set Specification
Geometry
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Identities
Sets
Names
Letters
Objects