Why

A natural distribution to associate with a dataset is to assign to each outcome a probability which reflects the number of times it appears in the dataset.

Definition

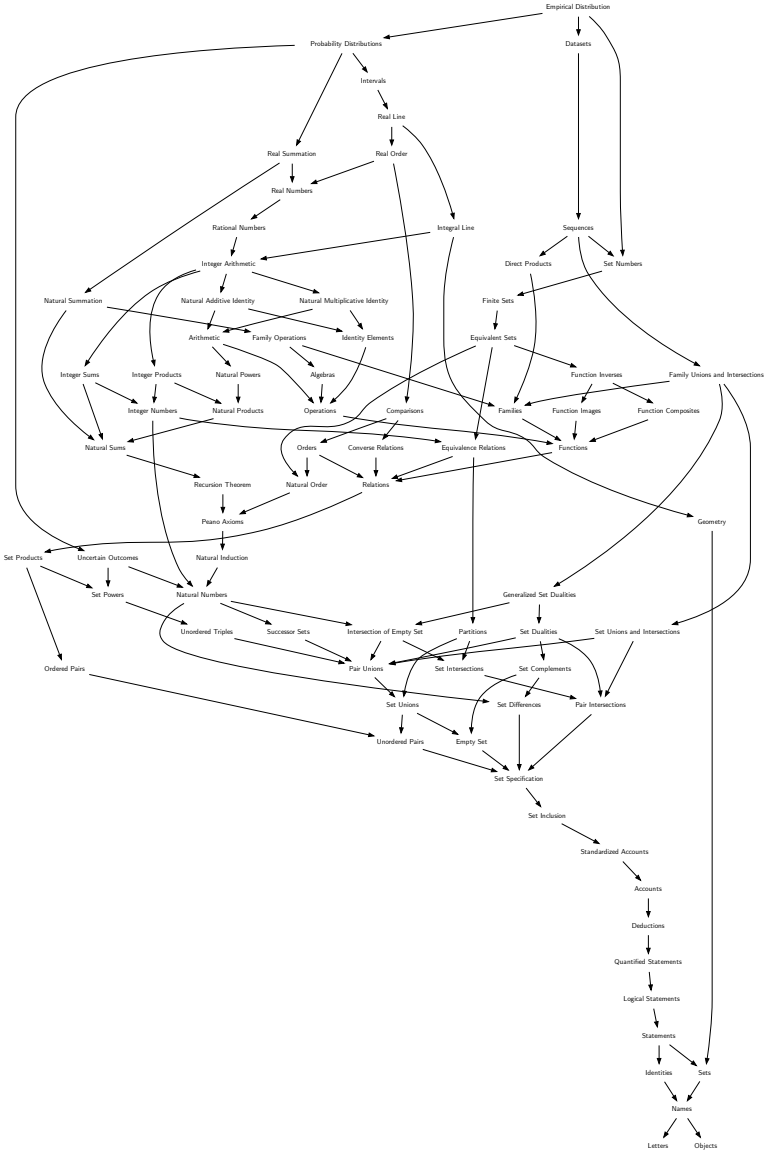
The *empirical distribution* of a dataset is the function which associates to each outcome the proportion times it appears in the dataset. Since the proporitions are nonnegative and sum to one, the function is a probability distribution.

Notation

Let A be a non-empty set and $(a^1, \ldots, a^n)$ be a dataset in A. The empirical distribution $q : A \to \mathbf{R}$ of $(a^1, \ldots, a^n)$ is defined by

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

In other words, to give the probability of outcome $a \in A$, we count the number of times it appeared in the dataset of size n, and then divide by n.


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