Why

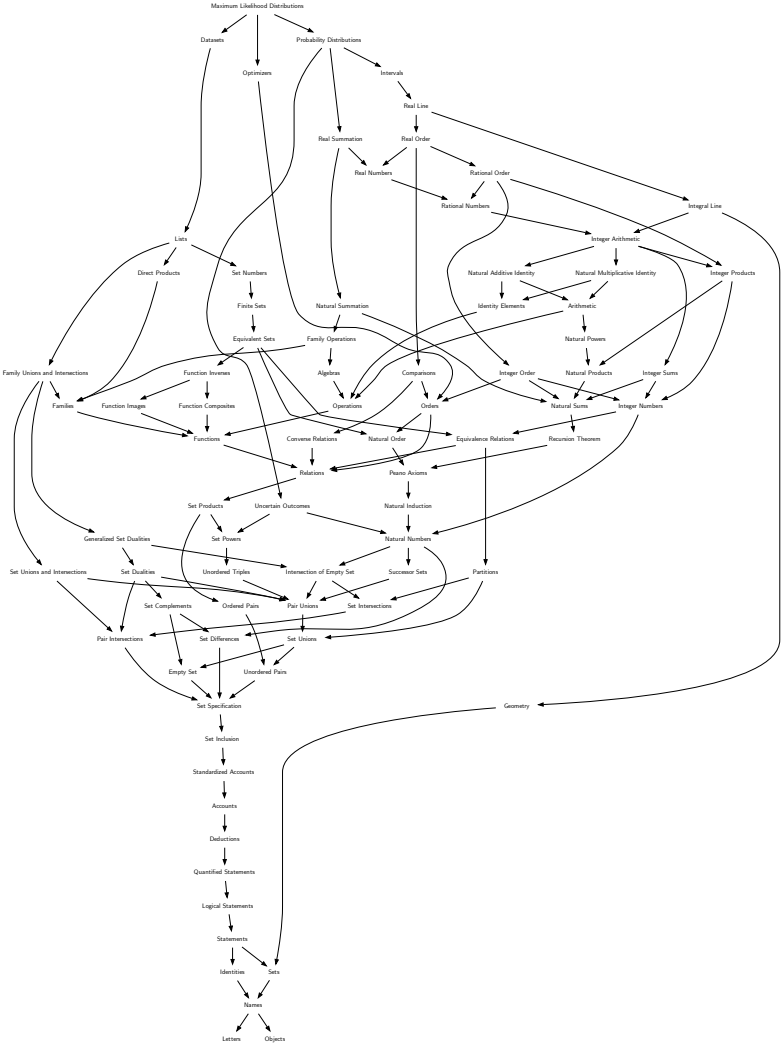
We want to summarize a dataset with a probability distribution.

Overview

Let $u^1, \ldots, u^n$ be a dataset in $\mathcal{U}$. We want a correspondence between the dataset and probability distributions over the set of outcomes A.

The *likelihood* of the dataset $u^1, \ldots, u^n$ is $\prod_{i=1}^n p(u^i)$. A *maximum likelihood distribution* $p^\star : \mathcal{U} \to \mathsf{R}$ is one which maximizes the likelihood over all distributions on $\mathcal{U}$.

We call the correspondence between datasets and distributions the *maximum likelihood algorithm*. We say that we are selecting the distribution according to the *maximum likelihood principle*. In general, we call any function from datasets to to distributions a *distribution selector*.


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