Why

We want to summarize a dataset with a distribution.

Overview

The *likelihood* (or *distribution likelihood*) of a probability distribution $p : A \rightarrow \mathbf{R}$ on a dataset $a^1, \ldots, a^n \in A$ is $\prod_{i=1}^{n} p(a^i)$. A *maximum likelihood distribution* $p^\star : A \rightarrow \mathbf{R}$ is one which maximizes the likelihood over all distributions on A.

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 distributions a *distribution selector*.

Maximum Likelihood Distributions

Probability Distributions

Optimizers

Intervals

Real Line

Real Summation   Real Order

Real Numbers   Rational Order

Rational Numbers   Integral Line

Integer Arithmetic

Lists

Set Numbers   Natural Multiplicative Identity   Natural Additive Identity   Integer Products

Direct Products   Finite Sets

Identity Elements   Arithmetic

Equivalent Sets   Natural Summation   Natural Powers

Family Unions and Intersections   Function Inverses   Family Operations   Comparisons   Integer Order   Integer Sums   Natural Products

Families   Function Composites   Function Images   Operations   Orders   Integer Numbers   Natural Sums

Functions   Converse Relations   Equivalence Relations   Natural Order   Recursion Theorem

Relations   Peano Axioms

Set Products   Natural Induction   Uncertain Outcomes

Generalized Set Dualities   Set Powers   Natural Numbers

Set Unions and Intersections   Partitions   Intersection of Empty Set   Set Dualities   Unordered Triples   Successor Sets   Set Symmetric Differences

Ordered Pairs   Set Intersections   Pair Unions   Set Complements

Pair Intersections   Set Unions   Set Differences

Unordered Pairs   Empty Set

Set Specification   Geometry

Set Equality

Set Inclusion

Deductions

Quantified Statements

Logical Statements

Statements

Identities   Sets

Names

Letters   Objects