Why

We want to approximate a given distribution with one which factors according to a tree.

Definition

Given $q : A \to [0,1]$, we want to find a distribution p on A and tree T on $\{1,\dots,n\}$ to

$$\begin{aligned} \text{minimize} \quad & d_{kl}(q,p) \\ \text{subject to} \quad & p \text{ factors according to } T. \end{aligned}$$

where d_{kl} is the relative entropy as a criterion of approximation. We call such a distribution a *tree distribution approximator* (or *tree approximator*) and we call the tree the *approximator tree.*

Result

Proposition 1. *Let $A_1,\dots,A_n$ be finite non-empty sets. Define $A = \prod_{i=1}^n A_i$. Let $q : A \to [0,1]$ a distribution and T a tree on $\{1,\dots,n\}$. The distribution $p_T^* : A \to [0,1]$ defined by*

$$p_T^* = q_1 \prod_{i \neq 1} q_{i|\mathrm{pa}\, i}$$

minimizes the relative entropy with q among all distributions on A which factor according to T.

Proof. Let $p : A \to [0,1]$ be a distribution which factors according to T. First, express

$$p = p_1 \prod_{i \neq i} p_{i|\mathrm{pa}_i}$$

where pa_i is the parent of vertex i in T rooted at vertex 1 ($i = 2,\dots,n$).

Second, recall that the relative entropy of q with p is $H(q,p) - H(q)$. Since $H(q)$ does not depend on p, p is a minimizer of the relative of q with p if and only if p is a minimizer of $H(q,p)$.

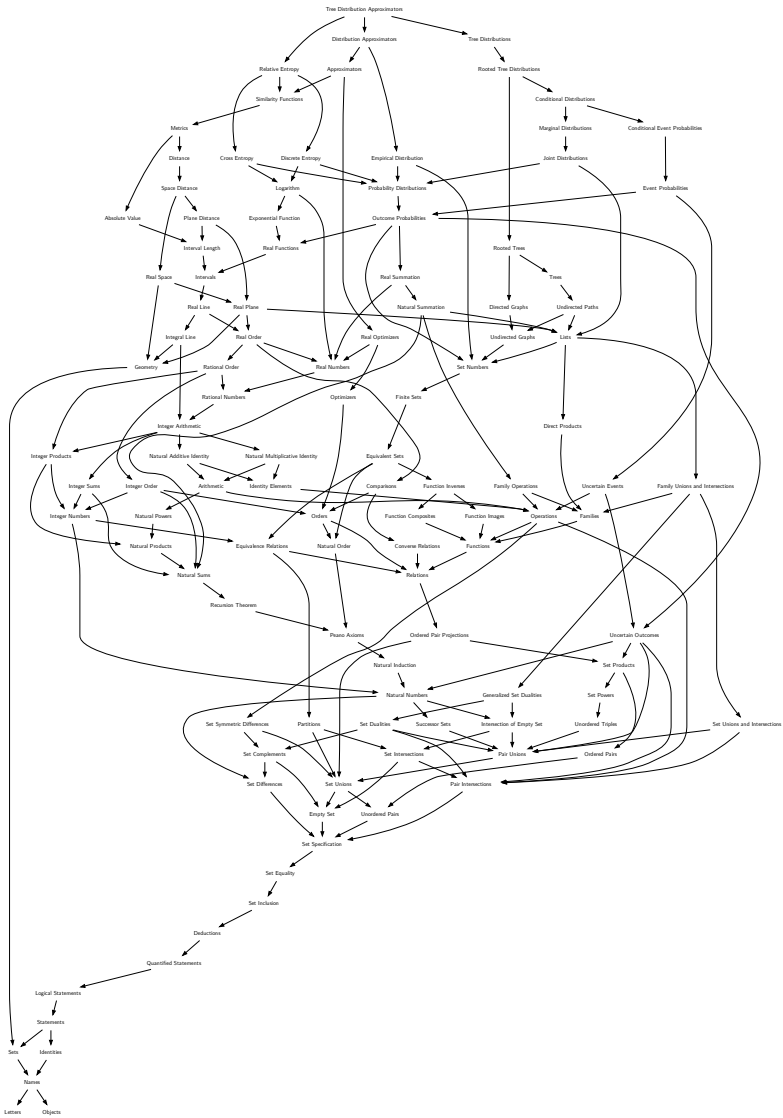
Third, express

$$
\begin{aligned}
H(q,p) &= -\sum_{a\in A} q(a)\log p(a) \\
&= -\sum_{a\in A} q(a)(\log p_1(a_1) + \sum_{i\neq 1} \log p_{i|\mathrm{pa}_i}(a_i, a_{\mathrm{pa}_i})) \\
&= H(q_1,p_1) + \sum_{i\neq 1} \sum_{a_{\mathrm{pa}_i}\in A_{\mathrm{pa}_i}} q_{\mathrm{pa}_i}(a_{\mathrm{pa}_i}) H(q_{i|\mathrm{pa}_i}(\cdot, a_{\mathrm{pa}_i}), p_{i|\mathrm{pa}_i}(\cdot, a_{\mathrm{pa}_i}))
\end{aligned}
$$

which separates across p_1 an $p_{i|\mathrm{pa}_i}(\cdot, a_{\mathrm{pa}_i})$ for $i = 2, \ldots, n$ and $a_{pa_i} \in A_{\mathrm{pa}_i}$.

Fourth, recall $H(\cdot,\cdot) \geq 0$ and is zero on repeated pairs. By this, we mean, for example, $H(p_1,p_1) = 0$. So $p_1 = q_1$ and $p_{i|\mathrm{pa}_i} = q_{i|\mathrm{pa}_i}$ are solutions. $\square$

The foregoing proposition states the form of an optimal approximator given a tree. A natural next question is to select the tree.


Tree Distribution Approximaton
Distribution Approximaton
Tree Distributions
Relative Entropy
Approximaton
Rooted Tree Distributions
Similarity Functions
Conditional Distributions
Marginal Distributions
Conditional Event Probabilities
Metrics
Distance
Cross Entropy
Discrete Entropy
Empirical Distribution
Joint Distributions
Space Distance
Logarithm
Probability Distributions
Event Probabilities
Absolute Value
Plane Distance
Exponential Function
Outcome Probabilities
Interval Length
Real Functions
Rooted Trees
Real Space
Intervals
Real Summation
Trees
Real Line
Real Plane
Natural Summation
Directed Graphs
Undirected Paths
Integral Line
Real Order
Real Optimizers
Undirected Graphs
Lists
Geometry
Rational Order
Real Numbers
Set Numbers
Rational Numbers
Optimizers
Finite Sets
Integer Arithmetic
Direct Products
Integer Products
Natural Additive Identity
Natural Multiplicative Identity
Equivalent Sets
Integer Sums
Integer Order
Arithmetic
Identity Elements
Comparisons
Function Inverses
Family Operations
Uncertain Events
Family Unions and Intersections
Integer Numbers
Natural Powers
Orders
Function Composites
Function Images
Operations
Families
Natural Products
Equivalence Relations
Natural Order
Converse Relations
Functions
Natural Sums
Relations
Recursion Theorem
Peano Axioms
Ordered Pair Projections
Uncertain Outcomes
Natural Induction
Set Products
Natural Numbers
Generalized Set Dualities
Set Powers
Set Symmetric Differences
Partitions
Set Dualities
Successor Sets
Intersection of Empty Set
Unordered Triples
Set Union and Intersections
Set Complements
Set Intersections
Pair Unions
Ordered Pairs
Set Differences
Set Unions
Pair Intersections
Empty Set
Unordered Pairs
Set Specification
Set Equality
Set Inclusion
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects