



Why

We want to summarize a dataset by a probability distribution.

Overview

We have a set of outcomes A . We have a dataset (a^1, \dots, a^n) in A . We want to define a correspondence between the dataset and probability distributions over the set of outcomes A .

We call correspondences *principles of distribution selection*.

Notice that we do not want a distribution on the record sequences. In other words, we want a distribution on A not a distribution on A^n .

