Why

We provide a principle for density selection based on maximizing the probability of the dataset.

Definition

Let $(x^1, \ldots, x^n)$ be a dataset in $\mathbf{R}^d$. The *density likelihood* for this dataset is

$$\prod_{k=1}^{n} f(x^k).$$

A *maximum likelihood density* is a density which maximizes the density likelihood among all densities. A *maximum likelihood density selector* is a density selector that associates a dataset with a maximum likelihood density.