



Why

We want to summarize a dataset in \mathbf{R} with a density

Definition

The *likelihood* (or *density likelihood*) of a density $f : \mathbf{R} \rightarrow \mathbf{R}$ on a dataset $x^1, \dots, x^n \in \mathbf{R}$ is $\prod_{k=1}^n f(x^k)$. A *maximum likelihood density* is a density which maximizes the likelihood among all densities.

As with probability distributions, we say that we are selecting the distribution according to the *maximum likelihood principle*. In general, we call any function from datasets to densities a *distribution selector*.

