



**Why**

We have a distribution to approximate. It is, for some reason, unsuitable to our needs and we want to replace it with one more suitable.

**Definition**

A *distribution approximator* is an approximator of a probability distribution. It is also a distribution. The criterion of approximation is any similarity function on distributions over the same space.

**Notation**

Let  $A$  be a non-empty set and  $q : A \rightarrow \mathbf{R}$  be a distribution. Let  $p : A \rightarrow \mathbf{R}$  be a distribution. Then  $p : A \rightarrow \mathbf{R}$  is a distribution approximator of  $q$ .

**Reasons for approximation**

**Infeasible to represent.** If there are many outcomes, many numbers are required to specify the distribution. If  $p : A^n \rightarrow \mathbf{R}$  where  $|A| = k$ , then there are  $k^n$  outcomes; take, for example,  $k = 2$  and  $n = 100$ . So we might want to find a distribution which requires fewer numbers to express. In other words, we want a different distribution, selected from the set of those which is easier to express, which is close to the original.

**Unreasonable from common sense.** The distribution

may be unreasonable as a result of our common sense. For example, it may give zero probability to an outcome which we know to be possible, and would like to model with non-zero probability. This may happen when working with an empirical distribution: a particular outcome does not appear in the dataset, however, our common sense suggests it is possible. In this case, we want to find a different distribution, selected from the set of those which is more reasonable based on common sense, which is close to the original.

