



Why

We want to talk about probability over finite sets.

Definition

A *probability distribution* or *probability mass function* is a real-valued function from a set of outcomes which is non-negative and normalized. A real-valued function on a finite set is *normalized* if the sum of its results is 1. We will refer to these as *distributions*. The *probability of an outcome* is the result of the outcome under the distribution.

Notation

Let A be a set of outcomes and let $p : A \rightarrow \mathbf{R}$ be a distribution. Then

$$p(a) \geq 0 \text{ for all } a \in A \text{ and } \sum_{a \in A} p(a) = 1.$$

PROPOSITION 1. *If $p : A \rightarrow \mathbf{R}$ is a distribution, then $p(A) \subset [0, 1]$.*

Proof. Let $a \in A$. First, $p(a) \geq 0$ by definition. Second, since p is normalized, $\sum_{b \in A} p(b) = 1$. And $p(a) \leq \sum_{b \in A} p(b)$, so $p(a) \leq 1$.

□

