

PROBABILITY DISTRIBUTIONS

Why

We talk about an uncertain outcome by assigning credibility to various outcomes. We use our intuition of proportion to do so. We start with 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.

The probability of an outcome is meant to indicate the credibility of the outcome. The word probability has its roots in the English word probable, which has the Middle English sense "worthy of belief". The probability of an outcome then is how worthy of belief it is. In the case of flipping a coin, or rolling a die, all outcomes are equally worthy of belief.

When an outcome has a higher probability we say that is is *more probable* than another outcome. Remember that the intuition is that we are saying the outcome is more "credible", or more "worthy of belief."

Notation

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

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

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

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

Example: coin

When flipping a coin, both a heads and a tails are worthy of belief. Thus we say that the outcome 0 (tails) has probability 1/2 and likewise for the outcome 1 (heads) In this case we say that we are modeling a *fair coin*.

Example: die

When rolling a die, all six sides are worthy of belief. Thus we say that the outcomes 1, 2, 3, 4, 5, 6 each have probability 1/6. Prior to roll, each is equally credible. In this case we say that we are modeing a *fair die*.

