Why

To talk about uncertain outcomes, we assign credibility to each outcome according to our intuition of proportion.[1]

Definition

A *probability distribution* (or *distribution*, *probability mass function*, *pmf*, *proportion distribution*) is a real-valued function from a finite 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 an outcome. The word probabiliy has its roots in the English word probable, which has the Middle English sense "worthy of belief". The probability of an outcome models how worthy of belief it is, relative to other outcomes. 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 than a second outcome we say that the first is *more probable* than the second. Similarly we say that the first outcome is *less probable* than the second outcome if it has higher probability than the second.

Notation

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

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

[1]Future editions may drop the dependence on real numbers, and use intuition of repeated trials to introduce *rational* probability distributions.

Simple consequences of definition

Proposition 1. *Suppose $p : A \to \mathbf{R}$ is a distribution. For all $a \in A$, $p(a) \leq 1$.*

Proof. Let $a \in A$. By definition, $p(a) \geq 0..$ Second, since p is normalized, $p(a) \leq \sum_{b \in A} p(b) = 1$. $\qquad \square$

Consequently, the range of p is contained in $[0, 1]$. For this reason, we will often introduce a distribution $p : A \to [0, 1]$.

Example: coin

When flipping a coin, both a heads and a tails are equally 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 equally 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.*

Probability Distributions

Intervals

Real Line

Real Order

Real Summation

Rational Order   Real Numbers

Integral Line

Rational Numbers

Integer Arithmetic

Natural Multiplicative Identity   Natural Additive Identity   Integer Products

Arithmetic   Identity Elements   Natural Summation

Integer Order   Integer Sums   Natural Powers   Family Operations

Comparisons   Integer Numbers   Natural Products   Families   Operations

Converse Relations   Orders   Equivalence Relations   Natural Sums   Functions

Relations   Natural Order   Recursion Theorem

Ordered Pair Projections   Peano Axioms

Set Products   Natural Induction   Uncertain Outcomes

Ordered Pairs   Set Powers   Natural Numbers

Partitions   Intersection of Empty Set   Unordered Triples   Successor Sets

Set Intersections   Pair Unions   Set Symmetric Differences

Set Unions   Set Complements   Pair Intersections

Unordered Pairs   Empty Set   Set Differences

Geometry   Set Specification

Set Equality

Set Inclusion

Deductions

Quantified Statements

Logical Statements

Statements

Sets   Identities

Names

Letters   Objects