Why

If we model some measured value as a random variable with induced distribution $p : V \to \mathbf{R}$, then one interpretation of $p(v)$ for $v \in V$ is the *proportion* of times in a large number of trials that we *expect* to measure the value v.[1]

Definition

Given a distribution $p : \Omega \to \mathbf{R}$ and a *real-valued* outcome variable $x : \Omega \to \mathbf{R}$, the *expectation* (or *mean*) of x under p is $\sum_{\omega \in \Omega} p(\omega)x(\omega)$.

Notation

We denote the expectation of x under p by $\mathbf{E}_p(x)$. When there is no chance of ambiguity, we write $\mathbf{E}(x)$.

Properties

Let $x, y : \Omega \to \mathbf{R}$ be two outcome variables and $p : \Omega \to \mathbf{R}$ a distribution. Let $\alpha, \beta \in \mathbf{R}$. Define $z = \alpha x + \beta y$ by $z(\omega) = \alpha x(\omega) + \beta y(\omega)$. Then $\mathbf{E}(z) = \alpha\mathbf{E}(x) + \beta\mathbf{E}(z)$. Many authors refer to this property as the *linearity of expecation*.

Example: expectation

Suppose $\Omega = \{1, 2, 3, 4, 5\}$ with $p(1) = 0.1$, $p(2) = 0.15$, $p(3) = 0.1$, $p(5) = 0.25$ and $p(5) = 0.4$. Define $x : \Omega \to \mathbf{R}$ by

$$x(a) = \begin{cases} -1 & \text{if } a = 1 \text{ or } a = 2, \\ 1 & \text{if } a = 3 \text{ or } a = 4, \\ 2 & \text{if } a = 5. \end{cases}$$

[1] Future editions may modify this explanation, and take a genetic approach via summary statistics.

The expectation of x under p is

$$\mathbf{E}x = -1 - 0.15 + 0.1 + 0.25 + 2(0.4) = 0.9.$$

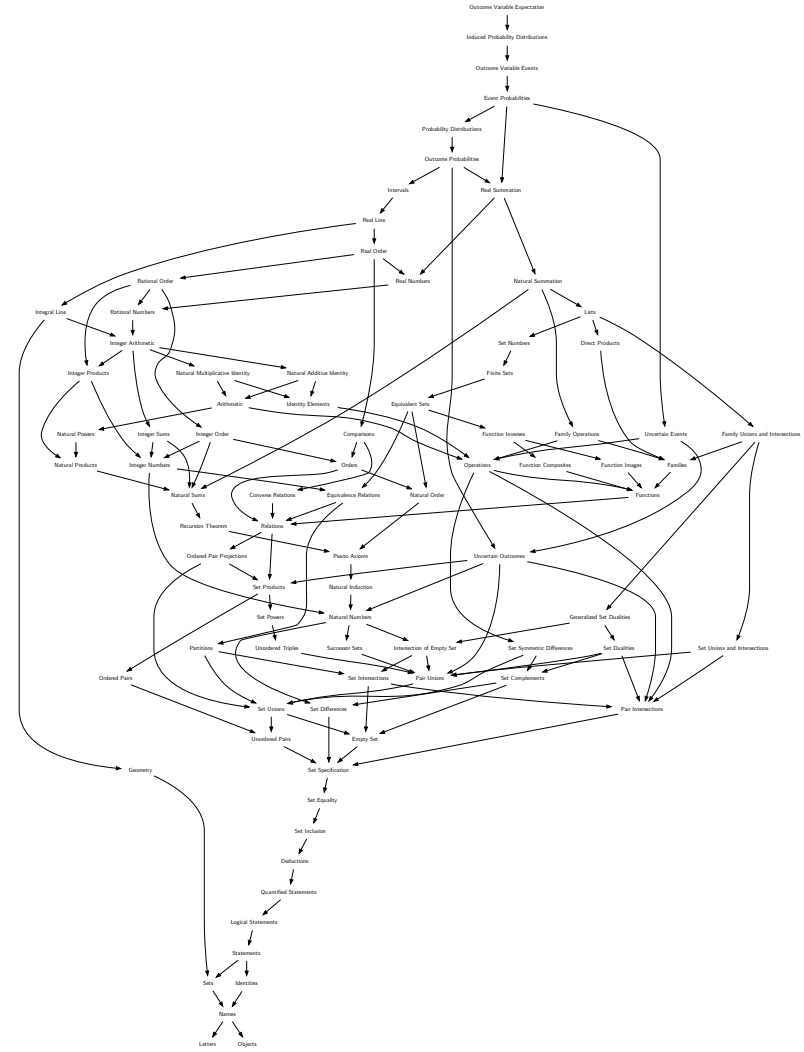
Two routes for computation

Denote by $p_x : V \to \mathbf{R}$ the induced distribution of $x : \Omega \to V$ (where $V \subset \mathbf{R}$). Then $\mathbf{E}(x) = \sum_{v \in V} p_x(v)v$ since

$$\begin{aligned}
\textstyle\sum_{\omega \in \Omega} p(\omega)x(\omega) &= \sum_{v \in V} \sum_{\omega \in x^{-1}(v)} x(\omega)p(\omega) \\
&= \sum_{v \in V} v \sum_{\omega \in x^{-1}(v)} p(\omega) \\
&= \sum_{v \in V} x(v)p_x(v).
\end{aligned}$$

Interpretations

We interpret the mean as the center of mass of the induced distribution.


Outcome Variable Expectation
Induced Probability Distributions
Outcome Variable Events
Event Probabilities
Probability Distributions
Outcome Probabilities
Intervals
Real Summation
Real Line
Real Order
Rational Order
Real Numbers
Natural Summation
Integral Line
Rational Numbers
Lists
Integer Arithmetic
Set Numbers
Direct Products
Finite Sets
Integer Products
Natural Multiplicative Identity
Natural Additive Identity
Identity Elements
Equivalent Sets
Arithmetic
Natural Powers
Integer Sums
Integer Order
Comparisons
Function Inverses
Family Operations
Uncertain Events
Family Unions and Intersections
Natural Products
Integer Numbers
Orders
Operations
Function Composites
Function Images
Families
Natural Sums
Convene Relations
Equivalence Relations
Natural Order
Functions
Recursion Theorem
Relations
Ordered Pair Projections
Peano Axioms
Uncertain Outcomes
Set Products
Natural Induction
Set Powers
Natural Numbers
Generalized Set Dualities
Partitions
Unordered Triples
Successor Sets
Intersection of Empty Set
Set Symmetric Differences
Set Dualities
Set Unions and Intersections
Ordered Pairs
Set Intersections
Pair Unions
Set Complements
Set Unions
Set Differences
Pair Intersections
Unordered Pairs
Empty Set
Geometry
Set Specification
Set Equality
Set Inclusion
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects