



Why

If we model some measured value as a random variable with induced distribution $p : V \rightarrow \mathbf{R}$, 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 .

Definition

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

Notation

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

Properties

Let $x, y : \Omega \rightarrow \mathbf{R}$ be two outcome variables and $p : \Omega \rightarrow \mathbf{R}$ a distribution. Let $\alpha, \beta \in \mathbf{R}$. Define $z = \alpha x + \beta y$ by $z(\omega) = \alpha x(\omega) + \beta y(\omega)$. Then $\mathbf{E}(z) = \alpha \mathbf{E}(x) + \beta \mathbf{E}(y)$.

