



Why

We can interpret the codomain of a random variable as a new sample space, since the underlying probability distribution induces a new probability distribution.

Definition

Let $p : \Omega \rightarrow \mathbf{R}$ be a probability distribution and $x : \Omega \rightarrow V$ an outcome variable. Define $q : V \rightarrow \mathbf{R}$ by

$$q(a) = \mathbf{P}[x = a].$$

Since events $x^{-1}(a)$ for $a \in V$ partition Ω , $\sum_{a \in A} q(a) = 1$. We call q the *induced distribution* (or *induced probability mass function*) of the random variable x . Thus we can think of V as a set of outcomes, which we call the outcomes *induced* by x .

Notation

It is common to denote it by p_x .

If $x : \Omega \rightarrow V$ is a random variable and $f : V \rightarrow U$, then if we define $y : \Omega \rightarrow U$ so that $y \equiv f(x)$, y is a random variable with induced distribution $p_y : U \rightarrow \mathbf{R}$ satisfying

$$p_y(b) = \sum_{a \in V | y(a)=b} p_x(a).$$

As a matter of practical computation, we can evaluate probabilities having to do with the outcome variable x using p_x instead of p . For example with x as in the example above, $\mathbf{P}(x = 4 \text{ or } x = 5) = p_x(4) + p_x(5)$, rather than $\sum_{\omega \in \Omega | x(\omega)=4 \text{ or } x(\omega)=5} p(\omega)$.

