



Why

Given a probability measure on the events of a set of outcomes, we can discuss probabilities of outcome variables.

Definition

Let $p : \Omega \rightarrow \mathbf{R}$ be a probability distribution with corresponding probability measure $\mathbf{P} : \mathcal{P}(\Omega) \rightarrow \mathbf{R}$. Suppose $x : \Omega \rightarrow V$ is an outcome variable. The *probability* $x = a$, for $a \in \Omega$, is

$$\mathbf{P}(\{\omega \in \Omega \mid x(\omega) = a\}).$$

From the definition of \mathbf{P} , we express the above as

$$\sum_{\omega \in \Omega \mid x(\omega) = a} p(\omega).$$

We refer to the probability of the *event* that $x = a$.

Notation

We denote the probability that $x = a$ by $\mathbf{P}[x = a]$. Our square brackets deviate from the slightly slippery but universally standard notation $\mathbf{P}(x = a)$. We prefer the square brackets, since $x = a$ is not itself an argument to \mathbf{P} , but shorthand for $(\{\omega \in \Omega \mid x(\omega) = a\})$.

Accordingly, there are many similar notations. For example $V = \mathbf{R}$, $\mathbf{P}[x \geq a]$ is $\mathbf{P}(\{\omega \in \Omega \mid x(\omega) \geq a\})$. Or, $\mathbf{P}[x \in C]$ means $\mathbf{P}(\{x \in \Omega \mid x(\omega) \in C\})$. Since the *event* that $x = a$ is the inverse image of $\{a\}$ under x , we also use the notation $x^{-1}(a)$. Or generally, $x^{-1}(C)$.

Example: sum of two dice

Define $\Omega = \{1, \dots, 6\}^2$ and define $p : \Omega \rightarrow \mathbf{R}$ with $p(\omega) = 1/36$ for each $\omega \in \Omega$. Define $x : \Omega \rightarrow \mathbf{N}$ by $x(\omega_1, \omega_2) = \omega_1 + \omega_2$. Then $\mathbf{P}[x = 4] = p((2, 2)) + p(1, 3) + p(3, 1) = 1/12$.

Induced probability

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 . It is common to denote it by p_x . Thus we can think of V as a set of outcomes, which we call the outcomes *induced* by 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 $\sum_{\omega \in \Omega | x(\omega)=4 \text{ or } x(\omega)=5} p(\omega)$.

