What is a probability?

The CHANCE Project

Version dated 4 April 2010
GNU FDL*

Definition 1 Suppose we have an experiment whose outcome depends on chance. The *sample space* of the experiment is the set of all possible outcomes. □

We first consider chance experiments with a finite sample space $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$. (The letter '$\Omega$', which is the usual choice for the name of the sample space, is the capitalized version of 'ω', the Greek letter 'omega'.) For example:

- We roll a die and the possible outcomes are 1, 2, 3, 4, 5, 6 corresponding to the side that turns up: $n = 6$ and
$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$
- We toss a coin and the possible outcomes are H (heads) and T (tails): $n = 2$ and
$$\Omega = \{H, T\}.$$
- We toss a coin three times: $n = 8$ and
$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$
- We roll a die twice: $n = 36$ and
$$\Omega = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), \dots, (6,6)\}.$$

It is frequently useful to be able to refer to some feature of the outcome of an experiment. For example, we might want to write the mathematical expression which gives the sum of four rolls of a die. For our sample space Ω we take the set of all 6^4 4-tuples (x_1, x_2, x_3, x_4) with $1 \le x_i \le 6$, $i = 1, 2, 3, 4$. On Ω we define four functions X_1, X_2, X_3, X_4 which report what turned up on each particular roll, so that, for example

$$X_1((2,1,2,5)) = 2;\; X_2((2,1,2,5)) = 1;\; X_3((2,1,2,5)) = 2;\; X_4((2,1,2,5)) = 5.$$

In symbols, $X_i((x_1, x_2, x_3, x_4)) = x_i$. Now if we define $S = X_1 + X_2 + X_3 + X_4$, S tells the sum of the four rolls. For example

$$\begin{aligned} S((2,1,2,5)) &= X_1((2,1,2,5)) + X_2((2,1,2,5)) + X_3((2,1,2,5)) + X_4((2,1,2,5)) \\ &= 2+1+2+5 \\ &= 10. \end{aligned}$$

Definition 2 A *random variable* is a function on the sample space. □

It has been said that 'a random variable is neither random nor a variable'. Indeed, according to our definition, a random variable X is a perfectly well-defined function. What is random and variable is the *value* of this function, namely $X(\omega)$, which depends on the particular outcome ω of our chance experiment. This confusion between a function and a particular value of the function is commonplace in mathematics. In algebra, we write $y = x^2$, and say that y is a 'dependent variable'. Change the value of the independent variable x and the dependent variable y changes along with it. In the same way, the value $X(\omega)$ depends upon the outcome ω, which is random. $X(\omega)$—and thus by abuse of language X—is a *random dependent variable.*

Note. In some cases the outcome of an experiment may be determined by the value of a single random variable. In fact, we can always arrange this: We simply take $X(\omega) = \omega$, the identity function, and then knowing the value of X tells us the outcome ω. In other cases, the outcome may be determined by the values of a set of random variables. In the dice-rolling example we've been discussing, knowing the values of the four random variables $X_1, \ldots, X_4$ (or of any four of the five random variables $X_1, \ldots, X_4, S$) determines completely the outcome of the experiment.

Now let us assign probabilities to the various possible outcomes of an experiment.

Definition 3 Let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be a finite sample space. A *distribution function* is a real-valued function m whose domain is Ω and which satisfies:

1. $m(\omega) \geq 0$, for all $\omega \in \Omega$, and
2. $\sum_{\omega \in \Omega} m(\omega) = 1$.

A *event* is any subset E of Ω. We define the *probability* of the event E to be the number $P(E)$ given by

$$P(E) = \sum_{\omega \in E} m(\omega) \ .$$

□

For example, let $\Omega = \{1, 2, 3, 4, 5, 6\}$ be the sample space which represents the roll of one die. A probability distribution m on Ω assigns to each $j = 1, \ldots, 6$ a nonnegative number $m(j)$ in such a way that

$$m(1) + m(2) + \cdots + m(6) = 1 \ .$$

For the case of a fair die we would assign equal probabilities or probabilities $1/6$ to each of the outcomes. For the event $E = \{1, 2, 3, 4\}$ we have

$$P(E) = \frac{4}{6} = \frac{2}{3}.$$

That is, the probability is $2/3$ that a roll of a die will have a value which does not exceed 4. A very common alternative way to express this is to take X to be the random variable telling the result of rolling the die (remember that formally, X is the identity function on Ω), and write

$$P(X \leq 4) = \frac{2}{3}.$$

Note that the argument of P here is not an event as such, but rather a formula that defines an event. When we write $P(X \leq E)$, what we really mean is the probability $P(E)$ of the subset E of Ω consisting of those ω for which $X(\omega) \leq 4$. As you can see, the abuse of language and notation is the probabilist's stock-in-trade!