

Two ways of determining an expectation value

A trial succeeds with probability p and fails with probability q ; $p+q=1$, and we confine our attention to $0 < p < 1$ (i.e. ignore the extreme cases in which a trial certainly succeeds or certainly fails).

An experiment is a sequence of trials up to and including the first trial that succeeds. The number of trials in an experiment is called the length of that experiment. We are asked to determine the expectation value of the experiment length, i.e. the average length, each length being weighted with the probability of the corresponding experiment.

In the following, k ranges over the positive integers. Since an experiment of length k consists of $k-1$ failing trials, followed by 1 trial that succeeds, and the successive trials are independent, probability theory tells us that

(0) the probability of an experiment of length k equals $q^{k-1} p$.

Hence, the expectation value x we were asked to determine is given by

$$(1) \quad x = (p + 2qp + 3q^2p + \dots)$$

and we are left with the duty to sum this series. To this end we observe

$$\begin{aligned} x &= (1 + 2q + 3q^2 + \dots)p \\ qx &= (q + 2q^2 + 3q^3 + \dots)p \end{aligned}$$

and find by subtracting

$$(2) \quad (1-q)x = (1 + q + q^2 + \dots)p$$

There are now three ways in which we can proceed.

- we know the sum of the geometric series, i.e. $(1 + q + q^2 + \dots) = 1/(1-q)$
- we don't know the above, and repeat the trick above, observing for y given by

$$\begin{aligned} y &= (1 + q + q^2 + \dots) \\ qy &= (q + q^2 + q^3 + \dots) \end{aligned}$$

by subtracting the two

$$(1-q)y = 1$$

- we interpret the right-hand side of (2):

$$p + qp + q^2p + \dots$$

in view of (0) as the probability that an experiment terminates. Because $0 < p$, i.e. $q < 1$, the probability of the experiment not terminating equals 0, hence the right-hand side of (2) equals 1.

Because of $1-q = p$, we find in all three cases $x = 1/p$.

This result is so simple that there must be a simpler way to derive it. There is. The equation

$$x = 1 + qx$$

follows because an experiment consists of 1 trial to begin with, followed with probability q by a new experiment.

The moral of the story is that there advantages in being a computing scientist, more precisely in remembering at the right moment that potentially infinite objects are most appropriately given recursive definitions. (Remember SASL's definition of an infinite sequence ONES of ones: $\text{ONES} = 1 : \text{ONES}$.)

Austin, 4 February 1991

prof. dr. Edsger W. Dijkstra
Department of Computer Sciences
The University of Texas at Austin,
Austin, TX 78712-1188
USA