Probability of Being Chosen in Repeated Independent Trials—C.E. Mungan, Spring 2012

Suppose that the probability of randomly selecting any object from a group of numbered objects in a trial is p (where $0 < p < 1$). When selected, an object is not removed from the group and so it can be selected again in future trials. What is the probability that any particular object will get selected *at least once* during the course of n trials?

Here are three different ways to see that the answer is $\boxed{1 - (1-p)^n}$.

solution #1

The odds of NOT choosing an object in one trial is $1 - p$. Since each trial is independent, the odds of not choosing the object in two trials is $(1-p) \times (1-p)$. Continuing, it is clear that the odds of being spared in n trials is $(1-p)^n$, and 100% minus those odds must therefore be the probability of getting chosen one or more times. This solution accords with the two intuitively obvious cases of $n = 1$, for which the odds of being selected are $1 - (1-p) = p$, and the limit as $n \to \infty$, for which the odds of selection are 100%. For example, if $p = 25\%$ then the odds of getting picked at least once in two trials is 7/16.

solution #2

We can verify any particular case we like. For example, consider the just mentioned case of $p = 25\%$ and $n = 2$. For simplicity we will suppose there are only four objects in the group, numbered 1 through 4. (Alternatively there could be say 100 objects and we choose 25 in each trial.) If we choose object 1 in the first trial and object 2 in the second trial, let's denote that outcome as "12." Here then is an enumeration of all 16 possible outcomes: {11,12,13,14, 21,22,23,24, 31,32,33,34, 41,42,43,44}. In how many of these outcomes does any particular number (say the number "1") occur? In 7 of them, so the odds are 7/16, just as predicted at the end of solution #1.

solution #3

Here is another way to enumerate all the possibilities for the same case. It explains why one cannot directly find the solution by calculating p^n instead of $(1-p)^n$. The probability that a "1" say will get selected is the sum of:

- the probability $(1/4)(1/4) = 1/16$ that "1" gets chosen both times (that is what p^n tells us),
- the probability $(1/4)(3/4) = 3/16$ that "1" gets chosen in the first trial but not in the second trial, and
- the probability $(3/4)(1/4) = 3/16$ that "1" is not chosen in the first trial but is selected in the second trial.

Those three values add up to 7/16, the same solution again.

The answer presented here differs from Marilyn vos Savant's claim that it is p independent of n at http://www.parade.com/askmarilyn/2011/12/Sundays-Column-12-25-11.html. The correct answer for the case she considers of $p = 25\%$ and $n = 4$ is $175/256 = 68.4\%$ not 25%.