

AFTERNOTE 1
To
INTRODUCTION

Dramatizing the Power of Simulation

A problem in probability is at the heart of every problem in statistics. Problems in probability often can be very confusing to the intuition. And formulas often either do not aid the intuition, or lead to the wrong answer, or both. But simulation can often get you the right answer and also help you understand why it is correct. Let's dramatize the point with puzzles:

1. If a family has two children and one of them is a boy, what is the probability that the other will also be a boy?

Most people—even professional statisticians—often quickly answer “one half.” That is not correct.

The solution is not obvious even after the person has decided to tackle the problem by simulation. It is not unusual for a person to say: I flip one coin, and if it is a head (boy), I then flip another coin and see how often that will be a boy, also, and then actually flip the coin once and record the outcomes. But that experiment does not resemble the situation of interest. A proper modeling throws *two* coins, examines to see if there is a head on *either*, and then examines the *other*.

Or consider a two-digit column of random numbers in Table i-A1-1, using odd numbers for females and even numbers for males. The first fifty lines are sufficient to suggest the correct probability, and also to make clear the mechanism: Two-female pairs, a fourth of the cases, are excluded from the sample. And mixed pairs—which give a “no” answer—are two-thirds of the remaining pairs, whereas the only pairs that give “yes” answers—two males—are only a third of the remaining pairs. The point of presenting the puzzle is this: Simulation gets it right very quickly and easily, whereas the deductive method of mathematical logic can result in much confusion.

This puzzle illustrates the power of simulation. And it supports by analogy the general use of the resampling method in probability and statistics because it reveals the limitations of human deductive capacity, even among those who are mathematically adept.

Someone might wonder whether formal mathematics can help us with this problem. Formal (even though not formulaic) analysis can certainly provide an answer. We can use what is known as the “sample space” approach which reasons from first principles; here it consists of making a *list of the possibilities*, and examining the proportion of “successes” to “failures” in that list.

Table i-A1-1 “Is The Other A Boy?” With Random Digits
(considering only cases where at least one is a boy)

Random Digits	Is the Other a Boy?	Random Digits	Is the Other a Boy?
25	N	50	N
92	N	28	Y
05	N	03	N
27	N	09	N
41	N	52	N
08	Y	22	Y
76	N	41	N
50	N	34	N
01	N	48	Y
48	Y	68	Y
69	N	06	Y
24	Y	87	N
42	Y	68	Y
81	N	43	N
90	N	23	N
82	Y	58	N
22	Y	26	Y
41	N	76	N
26	Y	10	N
61	N	21	N
88	Y	48	Y
26	Y	98	N
76	N	87	N
22	Y	36	N
23	N	47	N

Even = Boy Odd = Girl Cases with 2 girls ignored
Results: Y = 17 N = 33

First we write down the equally-likely ways that two coins can fall:

$$H_1H_2 \quad H_1T_2 \quad T_1H_2 \quad T_1T_2$$

noting that it is very easy to make a mistake in writing this list; great mathematicians have made such mistakes in the past even with problems as easy as this one (as we will see with the example of D'Alembert in just a minute).

Now we notice that if we have observed at least one head, the list of possibilities for the ways the two coins fell shrinks to H_1H_2 H_1T_2 T_1H_2 .

And now we can see that in only one of three of these possibilities would the "other" coin be a head. So the probability we ask about is $1/3$.

The crucial question is: Does this formal approach make the problem harder or easier than the simulation approach? That depends on your mental makeup. In practice, it turns out that the formal method seems to lead to a higher rate of error if *everyone* employs it than does the simulation approach. Hence we emphasize the simulation approach here, to lead to the highest rate of success (and enjoyment). But it is best if everyone finds the mode that is best for him or her.

Here's a puzzle I call D'Alembert's Misery. Great mathematicians have blundered on even simple problems in probability because they attacked them with only logical tools. One famous example was D'Alembert, living in the 18th Century (1717-1783; story in Schuh, 1968, p. 165).

D'Alembert asked: What is the chance of throwing at least one head in two tosses of a coin? He reasoned that there are three cases: tail on both tosses, tail then head, head on the first toss (no second toss necessary). Of these three cases, two are successes, and therefore the probability sought is $2/3$, he concluded.

What's the answer?

1. Toss two coins, and record number of heads.
2. Repeat (1) a hundred times.
3. Count number of outcomes with one or two heads (or more easily, the number with two tails), and divide by 100 to find the probability sought.

Compare D'Alembert's conclusion to the result of your Monte Carlo experiment performed as above. Some other fascinating puzzles in probability are found in Chapter 6.