Book Reviews

Will You Be Alive 10 Years From Now?

Introduction

The author has written many books for the general reader on mathematical and physics topics, but this is the first I've read. He has written before on probability: 'Duelling Idiots and Other Probability Puzzlers' was published in 2000. This new book is a collection of 25 'Challenge Problems' and an introduction which provides a quick tour through some standard classic 'puzzles from the past'

I like this book. The author does not shy away from mathematics as it's needed: double integrals, sums of binomial coefficients, are scattered through the pages. The author thus expects a certain mathematical expertise from his readership. However, anybody with a year's tertiary mathematics under their belt could manage this, and even high school students unafraid of pushing themselves a bit would enjoy it.

Another attraction of the book is its emphasis on technology: in particular computer programs for Monte Carlo approaches to these problems. The author uses a sort of Matlab/pseudocode hybrid — the use of Matlab stemming from his background as an electrical engineer. However, these should be readable by anybody with programming experience.

Most chapters contain any of three parts: a description of the problem, a 'theoretical analysis' and a 'computer simulation'. Only eight of the 25 chapters have all three.

The contents

The introduction contains a bit of history, with a discussion of the gamblers ruin: given a sequence of games between two persons with probability p that A will win any game, what is the probability that A will win; that is, that player B will eventually have paid A all his money? It is just this sort of problem which is the basis of historical probability, and the misunderstanding of which ensures that casino games are always loaded in favour of the house. There is an example of one of the few times Isaac Newton was wrong, and a nice section on *intransitive dice*. Such dice (three or more) will have sides with different numbers from the usual 1 through 6; the example given has three dice (called A, B and C) having sides $(18, 9, 8, 7, 6, 5)$, $(17, 16, 15, 4, 3, 2)$ and $(14, 13, 12, 11, 10, 1)$ respectively. The
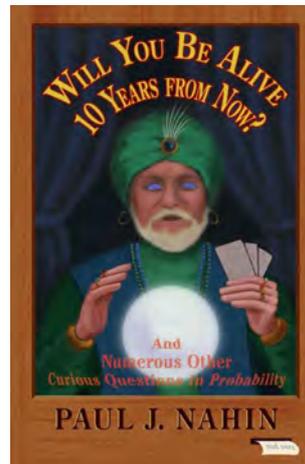
remarkable result is that in a sequence of throws (supposing the dice to be fair in the sense that each number is equally likely), then A will beat B with probability $21/36$; B will beat C with a probability of $21/36$, and C will beat A with a probability of $25/36$. Thus, A is more likely to beat B, which is more likely to beat C, which is more likely to beat A!

Prior to the introduction there is a Preface, where among his general remarks the author has a few digs at Marilyn vos Savant, whose (correct) analysis of the infamous Monty Hall problem caused a great stir at the time. However, she is not always correct, and in her analyses of some probability questions seems to be as likely to stumble as the rest of us.

One chapter considers the following problem: a class of 20 students, which includes two twins, is split randomly into five groups of four. What is the probability that the twins will be in the one group? There is a computer simulation, and a theoretical analysis which carefully counts all the ways in which 20 students can be split into five groups of four, and then all the ways in which the twins will be together. A nice example of the use of binomial coefficients. However, as so often happens, insight trumps brute force. Having given his analysis, the author then describes a solution provided by an early reviewer of the book, which takes about three lines of explanation, and no mathematics at all (other than one simple fraction). I imagine this was one of those 'why on earth didn't I think of that?' moments.

The most technical problem looks at differences between sides of a square; for example, if two points are chosen at random on two adjacent sides of a square, which is the average distance between them? This problem is harder to solve than it sounds, and the integrals used — the author shows solutions with both single and double integrals — become messy and complicated. However the author powers through them to obtain a closed-form expression which is then shown to be equal (or approximately equal) to the result obtained through simulation.

Of the problems given without a computer simulation, one of the most interesting is that of blood testing. Suppose you want to test a large group of people for the presence of a disease. Given the expense and difficulty of a blood test, you want to minimize the number of tests. One way is to pool the blood from several people into one batch and test that. If the result is negative, you know that all those people are free from the disease. If there is a low probability p that any person will be infected (the example given is testing soldiers in Word War II for syphilis, with $0.001 < p < 0.01$) then with k people, how do you manage the testing to minimize the number of tests? The analysis is in fact not difficult, but is explained with the clarity which characterizes the book.

There are also problems without theoretical analysis, such as determining the fairness of the ancient Jewish game 'dreidel': people put money, or sweets, or nuts, into a pot, and a four-sided spinner is spun by each person in turn. Depending on which face is uppermost, that person can either take everything from the pot, half of what's in the pot, nothing, or have to put one object in. Classic dreidel takes 'half' to mean the ceiling of $n/2$, given the discrete objects in the pot. The author curiously doesn't do this, but assumes that 'half' means 'divided by 2.0',

so that the value of the pot has arbitrary precision. It turns out to be unfair, in that whoever gets the pot (or half) early on can rarely be beaten after that, so that the game is loaded in favour of early spinners: the longer you have to wait to spin, the more unlucky you're likely to be.

As for the eponymous problem, the author shows how to calculate that probability from published Life Expectancy tables. This requires a fair dollop of calculus, and even an integral equation, which in turns requires the computation of a integral, which the author does by using Simpson's rule. If you want to know the answer for yourself, and if you're 50 or over, head off to the Department of Social Services website, and apply the author's reasoning.

Comments on the computer simulations

Although I like these very much, I think they are not optimal for understanding. It may have been better to stick with proper pseudo-code, as for example in *Introduction to Algorithms* by Cormen *et al.* [1]. The programs are all typeset in one font, with no distinction between key words and variables, which makes them hard to read. The programs are mostly in a sort of pseudo-Matlab, using loops instead of Matlab's far faster vectorization. The author defends this approach in his introduction when he claims (rightly) that more people would be familiar with loops than with vectorization. But in fact vectorization (which simply means that a function, applied to a matrix or vector, will be applied to every element) is extremely fast and efficient, as well as concise. If there had been a paragraph explaining this, then much of the book's code could be greatly simplified. For example, finding the average distance between points on two adjacent sides of the square takes only two lines in Matlab:

```
>> p = rand(1000000,2);
>> mean(sqrt(p(:,1).^2 + p(:,2).^2))
```

This might seem a little obtuse to people not used to Matlab, but I think if one is not going to use Matlab's powerful optimization, why not stick to a language like, say, Python?

There is a chapter called 'Bingo Befuddlement' which investigates a nontransitive Bingo. There are four cards:

$$A = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}, \qquad B = \begin{array}{|c|c|} \hline 2 & 4 \\ \hline 5 & 6 \\ \hline \end{array}, \qquad C = \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 4 & 5 \\ \hline \end{array}, \qquad D = \begin{array}{|c|c|} \hline 1 & 5 \\ \hline 2 & 6 \\ \hline \end{array}.$$

A single game consists of a permutation of 1–6, and the first player to fill a row wins. For example, with the game [3, 5, 6, 4, 2, 1], B wins by the time the 6 is called (B's second row: 5 and 6, have both been called); at that stage no other player will have been able to fill a row. The program given goes over two pages, and the explanation is even longer. But this can be vastly simplified with a little forethought: the only requirement is the index at which a given card wins. In Matlab, for a given permutation p, this takes only a few lines, the core of which are:

```
>> pt(p) = 1:6
>> min(max(pt(1),pt(2)),max(pt(3),pt(4)))
```

The first line gives the indices of 1–6 in the permutation p, the second line gives the index at which A will win. Similar lines for the other cards, wrapped in a loop for all permutations, and that's all it takes. You then find that each pair is tied for 72 permutations, and A beats B for 336 permutations, similarly for B beating C, C beating D, and D beating A. Thus:

$$\Pr(A \geq B) = \Pr(B \geq C) = \Pr(C \geq D) = \Pr(D \geq A) = 17/30.$$

What I'd like to know is, though: what are the rules by which such non-transitive Bingo cards can be created? And what is the upper limit on their number?

It turns out that this set of Bingo cards was designed by Donald Knuth, and explored by Martin Gardner [2]. There is no acknowledgment of this here, and in fact the lack of a bibliography is one of the few negative aspects of the book.

Who is this book for?

This is a difficult question to answer. The general reader, not mathematically trained, will find much of it impenetrable. In one sense, this is a pity, as the book contains much of general interest. So the book is designed for students, teachers, mathematicians, unafraid to get their hands dirty, and to explore some of the fascinations of probability.

Probability is one of those areas in which even experts can make mistakes. In answer to the question 'Why do people find probability unintuitive and difficult?', Sir David Spiegelhalter of Oxford has said: 'because *it is* unintuitive and difficult'. Humans are hard-wired to look for patterns, and to ascribe to the ones they imagine greater importance than those (possibly imaginary) patterns deserve. We are also notoriously bad at calculating risk [3]. I see Nahin's book as a great inducement to clear thinking, and we can all do with a good dose of that!

In spite of my gripes about the code and lack of bibliography, this is a wonderful book. The author writes with great clarity and vigour, and shows the hand of a master teacher. I thoroughly recommend it.

References

[1] Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C. (2009). *Introduction to Algorithms*, 3rd edn. MIT Press.

[2] Gardner, M. (1987). *Time Travel and Other Mathematical Bewilderments*. W. H. Freeman & Co.

[3] Schneier, B. (2006). *Beyond Fear: Thinking Sensibly About Security in an Uncertain World*, corrected edition. Copernicus.

Alasdair McAndrew
College of Engineering and Science, Victoria University, PO Box 14428, Melbourne, VIC 8001.
Email address: Alasdair.McAndrew@vu.edu.au