A parking problem

Peter J. Cameron

Have you ever wondered what your lecturers do all day when they are not giving lectures or classes or sitting in boring meetings? This is a story of a piece of mathematics, some of the details of which were worked out here at Queen Mary.

In the 1960s, the following problem was raised (originally in connection with data storage by computers).

> *A car park has n spaces in a line, numbered* $1,2,\ldots,n$*. The drivers of n cars have each independently decided on the position where they want to park. As each driver arrives at the car park, he drives to his preferred parking place. If the space is free, he parks there. If not, he drives on and takes the first available space; if he doesn't find an empty space, he leaves in disgust.*
>
> *What is the probability that all drivers manage to park?*

If $n = 2$, the probability is $\frac{3}{4}$, since only if both drivers choose 2 will they fail to park.

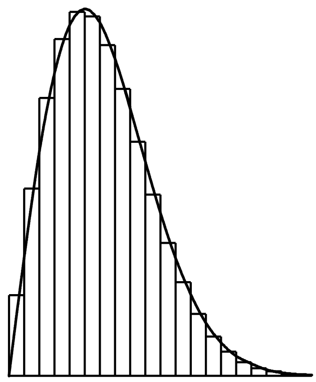
The answer is surprisingly simple, and there is a beautiful argument to show it. The required probability is $\frac{(n+1)^{n-1}}{n^n}$. There is not space in this article to give the proof (which was found by Henry Pollak of Bell Labs). This led to many extensions and generalisations involving algebra and combinatorics as well as computer science.

A year or two ago, I wondered about the number k of drivers who are unable to park. This number is a random variable: can its distribution be calculated? I calculated it for some small values, and Emil Vaughan (a PhD student in the department) pushed the calculations further, more-or-less by trying all possibilities. He was able to find the distribution up to $n = 20$.

At that point, I put the problem on my web page, so that all the world could have a go. The challenge was taken up by two computer science students from Saarbrücken in Germany, Daniel Johannsen and Pascal Schweitzer. They were able to find a recurrence relation for the numbers, and even more impressively, to solve it to give an explicit (though rather complicated!) formula.

I wanted to see a plot of the data. For this, the scale is very important. For example, since the limit of $(1+1/n)^n$ as $n \to \infty$ is e (the base of natural logarithms), we see that the probability that $k = 0$ (everybody parks) is about $\mathrm{e}/(n+1)$. If we scale the y-axis by a factor of n, this will be visible, but it turns out that some other values will be much too large to appear on the page.

After I gave a talk about the problem in the Friday Combinatorics Study Group, Dr Thomas Prellberg took it up. He found that the correct scaling was by the square root of n on the y-axis, and by $1/\sqrt{n}$ on the x-axis. If this is done, moreover, the histogram of the probabilities tends to a limit known as the *Rayleigh distribution*. I have shown a plot of the histogram for $n = 100$ (for $k > 20$, the probabilities are too small to appear on the graph), and on the comparable scale, the p.d.f. of the Rayleigh distribution. The curve has equation $y = 4x\mathrm{e}^{-2x^2}$.



In calculating this, Dr Prellberg had to evaluate a rather complicated integral. He was so pleased when he found the right substitution to do this that he set it as a prize question for the Calculus I class. So first-years struggling with the prize question: you can put some of the blame on me!