Mathematical Spectrum

Volume 13 1980/81 *Number 2*

A Magazine of *Contemporary Mathematics*

Published by the *Applied Probability Trust*

Mathematical Spectrum is a magazine for the instruction and entertainment of student mathematicians in schools, colleges and universities, as well as the general reader interested in mathematics. It is published by the Applied Probability Trust, a non-profit making organisation established in 1963 with the support of the London Mathematical Society. The object of the Trust is the encouragement of study and research in the mathematical sciences.

Volume 13 of *Mathematical Spectrum* will consist of three issues, of which this is the second. The first issue was published in September 1980 and the third will appear in May 1981.

Articles published in *Mathematical Spectrum* deal with the entire range of mathematical disciplines (pure mathematics, applied mathematics, statistics, operational research, computing science, numerical analysis, biomathematics). Both expository and historical material may be included, as well as elementary research and information on educational opportunities and careers in mathematics. There is also a section devoted to problems.

Articles are normally commissioned by the Editors; the Editorial Committee also welcomes the submission of suitable material, including correspondence, queries and solutions to problems, for publication in *Mathematical Spectrum*. All correspondence about the contents should be sent to:

The Editor, *Mathematical Spectrum*,
Hicks Building, The University, Sheffield S3 7RH.

Mathematical Spectrum Awards for Volume 12

Each year two prizes are available to contributors who are still at school or are students in colleges or universities. A prize of £20 is for an article published in the magazine and another £10 is for a letter or the solution of a problem. There were no articles in Volume 12 by authors eligible for the £20 prize. However, the editors have decided to award a prize of £10 to Gary Slater for his solution to Problem 11.8 (Volume 12, page 62) and a further prize of £10 to A. J. Granville for his letter *Tests for divisibility* (Volume 12, pages 93–94). We look forward to further contributions from readers.

The Series Swindle

JOHN PYM
University of Sheffield

The author is a Professor of Pure Mathematics at the University of Sheffield. He is a small man with glasses who always uses a 2B pencil when doing arithmetic because it rubs out easily. (Well, if Alfred Hitchcock could give himself small parts in his own films . . . .)

The heavy iron ball on the end of its chain followed the jib of the crane round in a long slow lazy arc. It disappeared through the wall of an old derelict house and for a moment all was still under the hot mid-day sun. Then the wall crumbled and the scene disappeared behind a thick haze of dust and rubble, leaving visible only the maniac grin of the demolition man perched in his cab high above the ruin around him.

'Nice day, Frank,' I said as I wandered into the only intact building in the desolate area. Frank Whitfield turned and spat. It was his way of being friendly; if he didn't like you, he didn't turn. He'd lived in that house, man and boy, for seventy-three years and by god he was going to die in it too, spite of all those councils, demolitions and that. When he'd reckoned he was too old to climb to the third floor, he'd rented it to me as an office. When he'd got too old to climb to the second floor, he'd rented it to me as a flat. Now he was keeping an eye on that fool in the crane: he didn't trust him. Nor would I have.

I'd just sat down at my desk when there was a soft knock at the door. I shouted 'Come in' as I slid the whisky bottle back in the drawer (I preferred to lunch in my office). In walked the kind of blonde men hoped was dumb because anything she said was bound to be less promising than the wiggle of her hips. She considered my battered furniture and decayed decor. 'You don't earn much. Aren't you any good?'

'Never failed yet,' I replied with an honest grin, 'but the kind of clients I get never seem to have enough money to pay.'

'Uh huh,' she said cynically. 'The private dick with the heart of gold.'

'If it was,' I said, 'I'd sell it. Come on, what's the problem?'

She sat on the box I used as a chair for clients. 'I've got this series,' she said. 'One third, one fifth, one seventh—'

'Dammit!' I interrupted. 'Not a next-term job!'

'Think I'm a fool?' she said contemptuously. 'I want the sum. What do you get if you add 'em all up?'

I put my feet on the desk, leaned back and eyed her coolly. I put her at about seventeen. 'Look, kid,' I said, 'teacher ain't gonna love you any more for getting a hard sum right.'

She was unperturbed. 'You stink,' she said emphatically and took a wad of papers out of her bag. She flipped through them as she talked. They were all tens. 'You're scum. But I need an answer by this time tomorrow. I'm stuck with you, god help me.' She put the money away and got up.

'OK,' I said. 'Where do I bring the answer?'

'I'll come back,' she said.

There was a dull thud and the floor shook. A few bricks fell out of the wall onto the floor. Through the new hole, a heavy iron ball could be seen floating away, and a distant jeering voice called 'Sorry, accident!' in unbelievable tones.

I said to her calmly, 'I may not be here tomorrow.'

I finished my lunch quickly and went down to my car. It was a long low job; the bonnet had large bulges in it showing the engine was much bigger than the space it was supposed to fit into. I looked up to where the crane-driver sat whistling in his box, slowly swinging his chain to and fro. Half an hour ago, my car had been a family saloon, but it now had the remains of a wall on top of it. I ran after a passing bus.

The address she'd given me was of a one-time mansion which had undergone an instant conversion job to turn it into a slum. I stopped at the door of what looked like a cupboard but bore the number of her flat and rang the bell loud enough to annoy the neighbours. All that happened was a piece of plaster fell from the moulded ceiling and a spider started to build its web in the gap. I rang the bell loud enough to drive the neighbours mad. All I roused was a fat tabby cat which waddled up and rubbed its back on my trousers. I wondered whether to kick it. 'Miaow' it said. Immediately a door across the landing opened.

'Treats that cat somethin' shockin', she does.' It looked like an old charlady.

'That's what we heard,' I said. This was no time to waste an opportunity. I guessed her sight wouldn't be too good so I got a miniature out of my pocket and waved it in front of her, concealing all but the label in my hand. 'My card,' I said. 'The name's John Walker. From the RSPCA. Can you tell me where I can find the young lady?'

She sniffed with a practised fruitiness. 'Lady? That'll be the day. She'll be at the bank.'

'We have to be polite,' I explained, but added ominously, 'in the beginning. Which one?'

My client was working as a cashier. The bank was L-shaped, on a street corner, so I was able to slip in, open an account with a couple of quid and fix an appointment to see the manager the next morning, all without her seeing me. Then I took off to the University.

I chose a door at random in the maths department, knocked and went in. Weedy little fellow, sitting behind a desk; glasses. He had a pencil in his hand and he looked as if he knew how to use it. In my trade, you learn not to underestimate such guys. 'What's your problem?' he asked.

'A third plus a fifth plus a seventh plus . . .,' I said. 'What's the total?'

'It doesn't converge,' he said.

'I just want to know what it all adds up to,' I explained patiently.

'I told you'—he sounded exasperated—'it doesn't converge. What class are you in?'

I strolled over and leaned on his desk. There was quiet menace in my voice. 'Look', I said. 'It's simple enough. Today I give you a third of a pound; tomorrow, a fifth; the next day, a seventh; and so on. How much cash do you end up with?'

'Sit down,' he said. As I slid into a chair, my hand moved onto my gun. His hand came up fast with a packet of chalk in it. He began to write on the board without a sound.

After 1 day: $\frac{1}{3}$

After 4 days: $\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{9}>\frac{1}{3}+\frac{1}{9}+\frac{1}{9}+\frac{1}{9}=\frac{1}{3}+3\,.\,\frac{1}{9}=2\,.\,\frac{1}{3}$

After 13 days:
$$\frac{1}{3}+\frac{1}{5}+\cdots+\frac{1}{27}=\frac{1}{3}+\left(\frac{1}{5}+\frac{1}{7}+\frac{1}{9}\right)+\left(\frac{1}{11}+\cdots+\frac{1}{27}\right)$$
$$>\frac{1}{3}+\left(\frac{1}{9}+\frac{1}{9}+\frac{1}{9}\right)+\left(\frac{1}{27}+\frac{1}{27}+\cdots+\frac{1}{27}\right)$$
$$=\frac{1}{3}+3\,.\,\frac{1}{9}+9\,.\,\frac{1}{27}=3\,.\,\frac{1}{3}$$

After $\frac{1}{2}(3^n - 1)$ days: $\frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{3^n}$

$$= \frac{1}{3} + \left(\frac{1}{5} + \frac{1}{7} + \frac{1}{9}\right) + \cdots + \left(\frac{1}{3^{n-1} + 2} + \frac{1}{3^{n-1} + 4} + \cdots + \frac{1}{3^n}\right)$$

$$> \frac{1}{3} + \left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9}\right) + \cdots + \left(\frac{1}{3^n} + \frac{1}{3^n} + \cdots + \frac{1}{3^n}\right)$$

$$= \frac{1}{3} + 3 \,.\, \frac{1}{9} + \cdots + 3^{n-1} \,.\, \frac{1}{3^n} = n \,.\, \frac{1}{3}.$$

I got it straight away. All you had to do was wait. When $n = 300$, you'd collected more than £100; when $n = 3000$, £1000; when $n = 3000000$—'Phew!' I said. 'You're saying the sky's the limit!'

He looked at me coldly. 'I'm saying there is no limit. What class are you in?'

I stood up. 'Thanks for the information,' I said, and threw a fiver onto his desk. As I went out, he began to laugh. I guess it was well above his usual fee, but what the heck.

The brochures which littered the bank manager's desk mentioned places like Las Vegas and Bangkok. Bodies goddesses might have envied cluttered golden beaches where a few bronzed heroes grinned at the thought of frequent indiscreet invitations. I guessed he was expecting some good fortune in which his wife and family were unlikely to share. 'Good morning.' I said. 'Going on holiday?'

Although it was only 9.30, it was already a blistering hot sunny day. I turned down the collar of my gaberdine mac. 'I need a loan,' I said. 'I'm thinking of thirty thousand. But the enterprise should give a big yield initially, then a decreasing one. I'd like to make repayments of a third of the sum after three months, a further fifth after six, a seventh after nine, and so on. I think this should eventually give you a reasonable interest on your money.'

I seemed to be looking out of the window, but I was concentrating on the pupils of his eyes. The movements are imperceptible so you have to use your imagination. I guessed this idea wasn't new to him. 'The bank isn't very ready to consider schemes which fall outside accepted practice,' he replied casually, 'but if there was adequate security I dare say I could arrange to lend you the money myself. There needn't be any rush to repay—the first instalment could come after six months if you preferred—but perhaps you could arrange to let me have the smaller sums more quickly. Could we say the second repayment after a further three months, the third after another one-and-a-half, and so on?'

'Sure,' I said quietly. I had nothing to lose as I wasn't going to do business with that crook. 'I'd prefer it. Get the thing over quicker.' You bet it would! The whole transaction would be completed in

$$6\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 6 \,.\, 1\Big/\left(1 - \frac{1}{2}\right) = 12$$

months. This guy was going to get rich quick.

'Perhaps I could offer you a sherry?' he said, going over to a cabinet.

'Whisky,' I said. Some compromises a man can't make.

The warmth in our relationship evaporated. 'The bank doesn't do business with men who drink whisky in the mornings,' he said.

When I got into the street, I spotted Charles Falconer walking in the bank's other entrance. The most elegant con-man in town doing business in the suburbs? I brooded. Then I ran and caught the bus to the University as it pulled away from the stop.

An hour later I slid unnoticed back into the manager's office. 'Don't press any emergency buttons,' I said, 'I only want you to look at this series.' I dumped my piece of paper on top of a picture of the sun-drenched grass skirts of Tahiti. He looked at it:

$$\sum_{n=1}^{\infty} \frac{1296}{1595n^6 - 8268n^5 + 10561n^4}.$$

'So what?' he said. 'Get out.'

'What's the first term?' I asked.

He couldn't resist it. He was one of those guys who attributes the decline of the nation to a lack of elementary arithmetic in schools. After a bit he said, 'One third.'

'The second?' I asked.

That was more of a strain. He got out a biro. '0·2', he said. 'Don't ask me another. Just get out.'

I looked at him with the calm superiority of a man who holds two aces his opponent doesn't yet know are in the pack. Then he got it. His voice shifted between a shrill panic and a hoarse despair. 'My god! A fifth! What's the third term?'

'One seventh,' I said with cold simplicity.

He banged his fists on the table in an infantile rage. 'Oh my god. What's the sum?'

Casually, I unbuttoned my mac. 'Well, I don't exactly know, but the way I figure it out is this.' I reached out and picked up a brochure. 'Mind if I write on this? I guess you won't be needing it any more.' He buried his face in his hands and began to sob. 'Pull yourself together man,' I said. 'This ain't too easy.'

'First thing,' I said, 'is to look at $p(x) = 1595x^2 - 8268x + 10561$. This is increasing if $dp/dx > 0$, that is $3190x - 8268 > 0$. So it's definitely increasing if $x > 3$. OK?'

He gave me a look like a lost sheep who'd been offered a street map. 'What do I lose?' he moaned.

'Dammit,' I said, 'be patient. I'm coming to that. Where we've got so far is that if $n \geq 4$, then $p(n) \geq p(4)$, and actually $p(4) = 3009$. It's also obvious that if $n \geq 4$, then $n^2 \geq 4^2 = 16$. Right?'

He'd gone back to whimpering. 'Come on!' I encouraged him. 'This is the interesting bit. For $n \geq 2$,

$$\frac{1}{n^2} < \frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n},$$

and so for any $N \geq 4$,

$$\sum_{n=4}^{N} \frac{1}{n^2} < \sum_{n=4}^{N} \frac{1}{n(n-1)} = \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) + \cdots + \left(\frac{1}{N-1} - \frac{1}{N}\right)$$
$$= \frac{1}{3} - \frac{1}{N} < \frac{1}{3},$$

get it?'

He didn't get it. He wasn't trying to get it. 'Just tell me the answer,' he pleaded. It's always like that. You try to explain the subtle points in the solution of a case and people just lose interest. I carried on relentlessly.

'So, if $N \geq 4$,

$$\sum_{n=4}^{N} \frac{1296}{1595n^6 - 8268n^5 + 10561n^4} = \sum_{n=4}^{N} \frac{1296}{p(n) \,.\, n^2 \,.\, n^2}$$
$$\leq \sum_{n=4}^{N} \frac{1296}{3009 \,.\, 16 \,.\, n^2} < \frac{1296}{3009 \,.\, 16 \,.\, 3} = \frac{9}{1003}.$$

Therefore

$$\sum_{n=1}^{N} \frac{1296}{1595n^6 - 8268n^5 + 10561n^4} < \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{9}{1003}$$
$$= \frac{72158}{105315} < 0{\cdot}69.$$

I contrived to write the last figure as if it came from the mouth of a dusky beauty lolling around in the Caribbean: we detectives thrive on cheap irony. 'However many instalments he pays, you can't get back more than 69 % of your money.'

'I'm ruined,' he said and began to cry in earnest. I let him suffer as I casually examined his office. I began to admire the man. Here was a guy who, by looking out of a window, could see how much rubbish came out of the back of one Chinese restaurant, and yet he was thinking of going East. I said abruptly, 'Stop the cheque.'

He was so far gone he didn't understand. But he looked up stupidly: 'Eh?'

'You guys inherit a certain kind of caution like elephants inherit trunks. You wouldn't dream of handling enough money to satisfy Charles Falconer in cash whatever method you were using to—shall we say, borrow?—it from the bank. Stop the cheque!'

He opened a drawer in his desk. I moved fast and pinned his hand down as it came into sight. It had half a bottle of whisky in it. He grinned at me with the pathetic gratitude of a dog when you stop kicking it. I prefer my cereal liquid; you don't get that damned crunching in the ears.

On my way out, I went up to a till and asked to withdraw two pounds from my account. She was startled. She demanded, 'How did you know where to find me?'

'I guess you wouldn't be so stupid as to hire a private dick who couldn't even discover where you worked,' I said bluntly.

'OK,' she said. 'What's my answer?'

'It looks good,' I said. 'He gets the lot. All the money in the world and then some!' She began to grin. 'In theory,' I added.

She turned hard. 'What d'you mean?'

'If I was thinking of blackmailing a chap involved in an underhand deal,' I explained without haste, 'I'd start by trying to find out who was going to come out on top. And if I discovered the two parties were Charles Falconer and a suburban bank manager, I'd give up the idea. The last guy who tried to blackmail Charles ended up £10000 down and in jail—after three weeks. If you want my advice—'

'I don't,' she said angrily. 'I hired you to get facts, not to give your goddamned opinion!'

'Maybe you did,' I said. 'But my guess is you aren't going to pay for either and I know which I'd rather give free. My advice is, you should give up listening at keyholes. Now I'll take those two quid in ones.'

She pushed a couple of grubby notes at me. I pushed them back. 'Do I have to teach you your job too?' I asked. 'Ones. Little round things. They go 'clink' in the pocket.'

I found Charles Falconer in a corner of the best restaurant in town beginning his lunch with smoked salmon and a bottle of Chablis. 'And a bottle of whisky for my friend,' he called to the waiter. 'Greetings, Charles,'—I was surprised to hear my childhood name, but we had shared a desk at school—'sit down, but for goodness sake take off that deerstalker. They went out with Holmes. This is a good day for me. I'm coming into some money.'

I waved the whisky away though I needed it. I said sadly, 'I'm afraid not, Charles.'

'Charles!' His tone was mock-severe. 'We're not working on opposite sides of the same case again are we? I thought I had a foolproof wheeze this time.'

'It was smart,' I said, 'but there was a flaw. You can't twist the meaning of words to suit yourself, whatever Humpty Dumpty might have thought. If a bloke standing at the bottom of a tower block reading a notice saying *Beware of the dog* is flattened by a St Bernard dropped from the twentieth floor, there's no judge in the land would say he was given fair warning. The words just don't usually mean that. And maths isn't basically any different, you know. It's all words at bottom. So the next term of your series, Charles, is one ninth, whatever weird formula you may have dreamed up to give something different.'

Charles looked a little glum. 'Then I suppose I shall have—as we say in the trade—to do a bunk?'

I smiled. 'I guess that won't be necessary. I was there when your sucker tore up the agreement.'

'You're a real friend,' said Charles, 'so I do hate to trouble you further. I have just sustained an attack of serious financial embarrassment. Could you manage a loan of a few pounds until next Tuesday? I imagine we shall meet at the Social Security office as usual?'

'That's right,' I said. 'I did get a job, but it didn't pay.'

On my way home I spent the last of my cash on a couple of masonry trowels and a bag of cement. I found Frank standing beside a large pile of rubble throwing bricks at the crane while the driver jeered. 'Hold it, Frank,' I said, shoving a trowel in his hand. 'Don't chuck the house away. Maybe we can rebuild enough of it to sleep in before nightfall.'

The Birthday Problem and Biological Research

SUSAN R. WILSON

The Australian National University

The author is currently a Fellow in the Department of Statistics, The Australian National University, and is in charge of consulting in the Research School of Social Sciences and the Research School of Pacific Studies. After graduating from the University of Sydney she obtained a Ph.D. degree from the Australian National University for research on statistical problems in genetics. She also does research on biological models.

The birthday problem is an abstract form of various problems occurring in biology. The biological applications will be briefly described later. First the problem of *repeated birthdays* is outlined.

Suppose that there are r people present in a room. What is the probability that no two persons in it have the same birthday? Let us assume that each person in the room can have his or her birthday on any one of the 365 days in the year (ignoring the existence of leap years), and that each day of the year is equally likely to be the person's birthday (ignoring changes in birth rates throughout the year due to seasonal and other effects). Now selecting a birthday for each person is the same as selecting a number drawn randomly from an urn containing $n = 365$ balls, numbered 1 to 365. In other words this birthday problem is the same as the problem of drawing a sample of size r with replacement from an urn containing n balls numbered 1 to n, where at each draw we note the number and replace the ball in the urn. Let P_r denote the probability that there are no repetitions in the sample (that is, that all the numbers in the sample occur just once), then it can easily be shown, see Feller (reference 1), that

$$P_r = \frac{n(n-1)\dots(n-r+1)}{n^r}$$
$$= \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{r-1}{n}\right).$$

For, there are a total of n^r different samples of size r which can be drawn with replacement. The number of these samples satisfying our condition (of no repetitions in the sample) is $n(n-1)\dots(n-r+1)$, since the first draw can be any of the n numbers, the second any of the remaining $n-1$ numbers, and so on. Various values of P_r for $r = 5, 10, 20, 22, 23, 24, 30, 40, 50, 60$ persons are given in Table 1.

TABLE 1

In a room containing r persons, P_r is the probability (correct to 3 decimal places) that there are not two or more persons in the room with the same birthday and $Q_r = 1 - P_r$ is the probability that there are two or more persons with the same birthday.

r	P_r	Q_r	r	P_r	Q_r
5	0·973	0·027	24	0·462	0·538
10	0·883	0·117	30	0·294	0·706
20	0·589	0·411	40	0·109	0·891
22	0·524	0·476	50	0·030	0·970
23	0·493	0·507	60	0·006	0·994

How many people do you think there must be in a room for the probability that at least two of them will have the same birthday to be greater than $\frac{1}{2}$? Students who have been asked this question have given answers as high as 100 and even 700. From Table 1 we see that the answer is only 23.

Now in obtaining the answer to the repeated birthday problem we made several assumptions which were only approximately true. Suppose we now wished to be more exact, and in fact had 366 *different* values for the probability that a birthday falls on any one of 366 days (including 29 February). Let p_1 be the probability of a person having his/her birthday on January 1, p_2 on January 2, ..., p_{366} on December 31, so that $\sum_{i=1}^{366} p_i = 1$. The probability of observing in our sample (of size r) r_1 birthdays on January 1, r_2 on January 2, ..., r_{366} on December 31 respectively, where $\sum_{i=1}^{366} r_i = r$, is

$$r! \prod_{i=1}^{366} \frac{(p_i)^{r_i}}{r_i!}, \qquad (0! = 1)$$

by the standard multinomial distribution formula, which is an extension of the well-known binomial formula when there are only two possible events. No common birthday occurs when r different $\{r_i\}$ have value 1 and the remaining $(366 - r)$ have value 0. The sum over *all* such outcomes is the probability of two or more people not having the same birthday, and is

$$P(r, 366) = \sum_{\Omega} r! p_{i_1} p_{i_2} \cdots p_{i_r}$$

where $\Omega = \{(i_1, i_2, \ldots, i_r)\colon i_j \in (1, 2, \ldots, 366), j = 1, 2, \ldots, r \text{ and } i_1 \neq i_2 \neq \cdots \neq i_{366}\}$. In other words Ω represents the selection of all $366!/(366 - r)!r!$ subsets of r probabilities from the set $\{p_1, p_2, \ldots, p_{366}\}$.

Computing $P(r, 366)$ from this formula is not a practical undertaking for many problems, especially those to be discussed later. For example, for our repeated birthday problem, if there were, say, $r = 30$ people in the room, this formula would require summing $366!/30!336!$ or approximately 5×10^{44} terms. Even on a modern automatic computer this would take about 10^{30} years. Recently a method of calculating $P(r, n)$, where n stands for 366 in the above formula, has been found by Gail, Weiss, Mantel and O'Brien (reference 2). They showed that $P(j, m)$ satisfies the following recursion formula:

$$P(j, m) = P(j, m - 1) + j p_m P(j - 1, m - 1)$$

for $j = 0, 1, \ldots, r$ and $m = 0, 1, \ldots, n$. For, either no two of the j persons will have the same birthday from among $m - 1$ dates, or if $j - 1$ have no common birthday from among $m - 1$ dates, and p_m is the probability of the mth date, there are j ways of selecting which of the persons will have this date for a birthday. Starting with appropriate boundary conditions, namely $P(j, 0) = 0, j = 1, 2, \ldots, r$ and $P(0, m) = 1, m = 0, 1, 2, \ldots, n$, one can use a high-speed computer to calculate probabilities like $P(30, 366)$ using this recursion. This calculation is relatively quick; it now takes less than one minute of DEC-10 central processor time.

Gail et al. (reference 2) were motivated to find a method of calculating $P(r, n)$ by the following biological problem. They had 8 enzyme loci; the first such locus had 3 allozyme genotypes (different 'categories'), the second 5 genotypes and so forth. There were thus $n = 3 \times 5 \times 3 \times 3 \times 3 \times 3 \times 5 \times 2 = 12150$ *distinct* allozyme signatures (i.e. vectors of 8 categories) with known non-zero probabilities. Too many common signatures suggest that the set of cell cultures has been contaminated by a single cell line. So one is interested in finding the probability that r randomly selected cell cultures will have no common signature. One striking result (for their data) was that the chances were greater than $\frac{1}{2}$ of a common signature if $r = 19$, despite the enormous number $n = 12150$ of signatures.

The above results are also applicable to other biological matching problems. One such problem is the determination of the probability of at least one good immunological match in a group of r individuals, where all immunotypes fall into one of n different categories within each of which the immunologic match is 'good'. Thus we see that the 'birthday problem' is of considerable practical relevance to modern research in biology.

References

1. W. Feller, *An Introduction to Probability Theory and its Applications*, Volume 1, second edition (Wiley, New York, 1957).
2. M. H. Gail, G. H. Weiss, N. Mantel and S. J. O'Brien, A solution to the generalized birthday problem with application to allozyme screening for cell culture contamination, *J. Appl. Prob.* **16** (1979), 242–251.

Mathematics and Sport

DAVID BURGHES
Cranfield Institute of Technology

Dr Burghes has lectured on mathematics at Sheffield and Newcastle Universities and presently teaches at Cranfield Institute of Technology, where his main research interests are in optimal control and in the dissemination of suitable teaching material in applied mathematics.

1. Introduction

The world to which mathematics is applied has been steadily growing this century. It seems that no area of human endeavour is immune from mathematical reasoning.

In this article we will give three illustrations of how mathematical analysis is used in running, shot and weight lifting. The models presented will be relatively simple ones; they will quite clearly not represent reality completely, but nevertheless they do have a practical value. They are used in competitions, by competitors and coaches in order to make decisions, or to improve techniques.

TABLE 1. World running records

Distance (m)	Time (s)	
	Men	Women
100	9·9	10·8
200	19·8	22·1
400	43·8	49·9
800	103·4	114·9
1000	113·9	—
1500	212·2	236·0
2000	291·4	—
3000	455·2	507·2
5000	792·9	—
10000	1650·5	—
20000	3444·2	—
25000	4456·8	—
30000	5490·4	—

2. Examples

(i) *Running*. In Table 1, we give the world records (as at December 1977) for both men and women for various distances. This data is taken from the IAAF publication *Progressive World Record Lists* (reference 1). The times, T, for men, are plotted against distance, D, in Figure 1. If we are looking for a relationship between T and D, there are two distinct ways of tackling the problem. The first is to look at a theoretical model for running (see Keller (reference 2)), which would require both dynamic equations of motion and also an oxygen balance equation. We would also have to distinguish carefully between short races, where the runner is accelerating for a significant amount of time, and longer races which are run at approximately

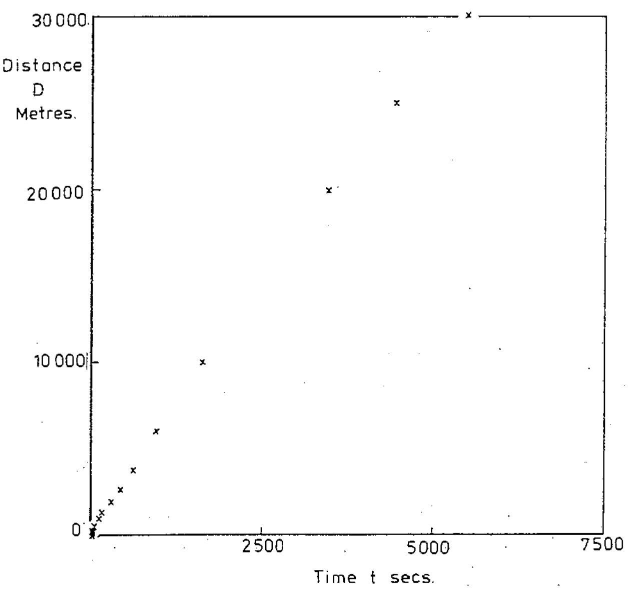


Figure 1. Distance, D, against time, T (for men).

constant speed. The second approach, and the one that we shall adopt, is to try to arrive at a model based not on theoretical arguments but purely on the available data (we usually call this an empirical model).

Returning to Figure 1, we first note that it is almost impossible to represent clearly the data for shorter distances and for longer races on one graph. One way out is to draw a log D – log T graph, and this is shown in Figure 2. We can now represent

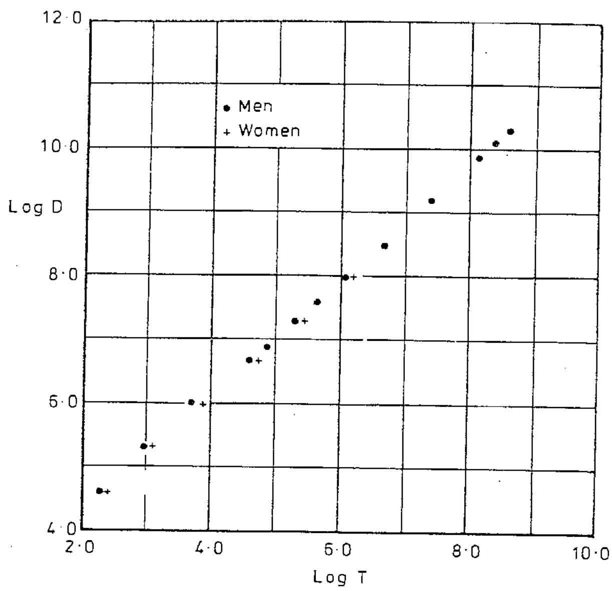


Figure 2. Log D against log T.

all the data points quite clearly, and it is not difficult to see that a straight line appears to be an extremely good fit for the data for both men and women. The slope of the straight line is approximately 0·9, so that we can write

$$\log D = (0{\cdot}9)\log T + \text{constant}$$

i.e.

$$D = kT^{0{\cdot}9};$$

or approximately

$$T = KD^{1{\cdot}1}. \qquad (1)$$

(Of course, we do need to check, using for example, a least-square analysis, that the straight line is really a good fit.)

This model implies that $T/D^{1{\cdot}1}$ should be constant, and so provides a method of handicapping different distance races. The value $T^* = T/D^{1{\cdot}1}$ is the handicapped time.

This analysis is used to

(a) determine a runner's 'best' distance to race over;
(b) compare runners over their chosen distances;
(c) find an overall running champion in competition (for example, using the data in Table 1, the 200 m world record holder, Don Quarrie of Jamaica, gives the least value of T^*).

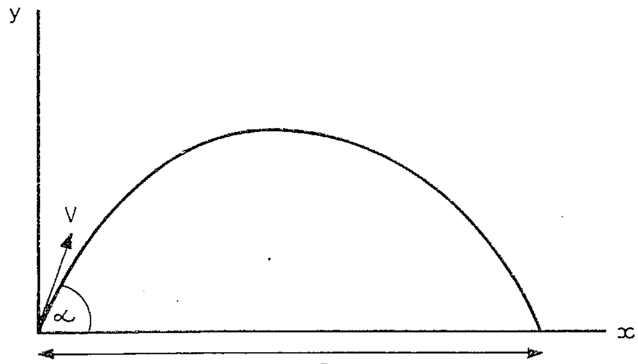


Figure 3. Shot trajectory.

(ii) *Shot.* The basic problems faced by a shot putter are how to accelerate across the circle in order to give maximum momentum to the shot, and at which angle to release the shot. In order to answer the second question, we start with a very simple particle model, as illustrated in Figure 3. We assume the motion is 2-dimensional and suppose that the shot leaves with speed v and angle α to the horizontal. Assuming the only force acting is constant gravity, the equations of motion ($\ddot{y} = -g, \ddot{x} = 0$) give a predicted range

$$R = 2V^2 \cos\alpha \sin\alpha/g \tag{2}$$

assuming $y = 0$ *at the initial release*. Clearly, for maximum R, we must choose $\alpha = \pi/4$; we also note that R depends on the *square* of v; so it is probably more important to aim at increasing v at the expense of an accurate angle α. For example, a 5 % increase in v yields a 10 % increase in R; whereas a 10 % error in the angle, away from $\pi/4$, only produces a 1 % reduction in R.

The assumption $y = 0$ initially, though, is not in good agreement with reality. If we assume $y = h$ initially, it is easy to show that the range is now given by

$$R = v\cos\alpha(v\sin\alpha + (v^2\sin^2\alpha + 2hg)^{1/2})/g. \tag{3}$$

To determine the value of α for maximum range, we must evaluate $dR/d\alpha = 0$. Now

$$gR/v^2 = \cos\alpha(\sin\alpha + (\sin^2\alpha + \beta)^{1/2})$$

where $\beta = 2hg/v^2$, and so $dR/d\alpha = 0$ yields

$$\sin\alpha(\sin\alpha + A^{1/2}) = \cos\alpha(\cos\alpha + A^{-1/2}\sin\alpha\cos\alpha)$$

where $A = \sin^2\alpha + \beta$. Thus

$$\cos 2\alpha = A^{1/2}\sin\alpha - A^{-1/2}\sin\alpha\cos^2\alpha$$

i.e.

$$A^{1/2}\cos 2\alpha = \sin\alpha(\beta - \cos 2\alpha).$$

Squaring and substituting for A eventually gives

$$\cos 2\alpha = \beta/(2 + \beta)$$

and so

$$\sin\alpha = 1/[2(1 + hg/v^2)]^{1/2}. \tag{4}$$

(Note that we can check this answer in the special case $h = 0$, giving $\sin\alpha = 1/\sqrt{2}$, as we saw above.)

For suitable values of h and v, (4) can now be used to deduce the theoretical optimum angle of projection. For example, if $h = 2\,\text{m}$, $v = 5\,\text{ms}^{-1}$, then $\alpha = 0{\cdot}56$ radians (32°).

Of course the model is still a very simple one, and can be criticised in many ways. An improved model can be obtained by using a 2-particle approach, the putter and the shot, and considering the impulse given to the shot. Interested readers might like to try this approach. It could also be of significance that the initial projection speed, v, depends, to a certain extent, on the projection angle, α!

TABLE 2. 1976 Olympic Games results for weightlifting

Bodyweight classes		Lifts (kg)		
Name	Maximum weight (kg)	Snatch	Jerk	Total
Flyweight	52	105·1	137·7	242·8
Bantamweight	56	117·7	145·2	262·9
Featherweight	60	125·2	160·2	285·4
Lightweight	67·5	135·2	172·8	308·0
Middleweight	75	145·2	190·3	335·5
Light-heavyweight	82·5	162·7	202·9	365·6
Middle-heavyweight	90	170·3	212·8	383·1
Heavyweight	110	175·3	225·4	400·7
Super-heavyweight		185·3	228·2	413·5

(iii) *Weightlifting*. Table 2 shows the 1976 Olympic Games Results for the two lifts, the snatch (where the weight is pulled from the floor to a locked arms overhead position in a single move) and the jerk (where two movements are allowed, the first to the chest and the second to the overhead position). There are nine bodyweight classes, and as expected, as bodyweight increases so does the lift.

The data is illustrated in Figure 4, on the assumption that the lifters are all at their maximum allowed bodyweight.

One of the difficulties in smaller competitions is that there are not enough competitors in each weight class. We therefore require a method of handicapping, so that lifters of any bodyweight can be compared. Such a form of handicapping could also be used to find an overall winner for the 1976 Olympic Games competition.

In (i), we developed a model directly from the data. In this example we will follow the alternative method and start by developing a simple theoretical model. We assume that the maximum weight, say L, which can be lifted by a weightlifter of bodyweight W is proportional to the cross-sectional area, say A, of the lifter's muscle; in mathematical terms, $L = k_1 A$. Secondly, we assume that A is

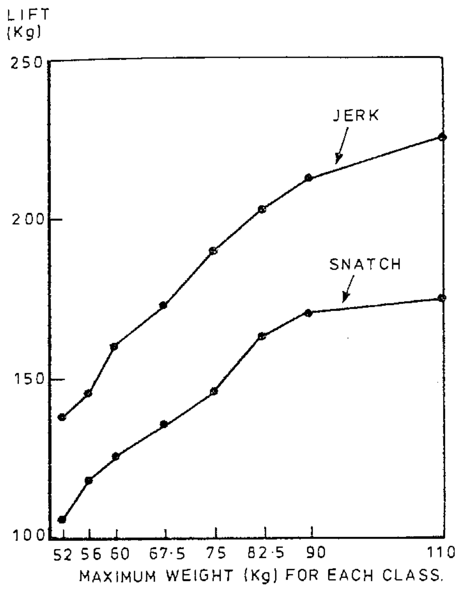


Figure 4. Snatch and jerk for bodyweight classes.

proportional to a typical body length, say l, squared: $A = k_2 l^2$. The body length is defined from $W = \rho l^3$ where ρ is the body density. Combining these three equations, we obtain

$$L = KW^{2/3}. \tag{5}$$

This model predicts that $L/W^{2/3}$ should be a constant. A quick check shows that this is not altogether true for the data above. On the other hand, the data is not truly representative; and indeed we would not expect the values of $L/W^{2/3}$ to be exactly constant. Using this model, the handicapped lift is $L' = L/W^{2/3}$, the overall winner being the one with the largest value of L'. For the data in Table 2, the overall winner is the light-heavyweight for both the snatch and the jerk. This particular model has been extensively used, particularly for junior and minor competitions. Many other models have also been used, notably:

(a) Austin formula: $L' = L/W^{3/4}$; (6)
(b) TV Superstars: $L' = L - W$; (7)
(c) O'Carroll formula: $L' = L/(W - 35)^{1/3}$. (8)

This last formula (see reference 3 for its derivation, which is partly theoretical and partly statistical) is the current method used for handicapping in serious competitions in the U.K.

Interested readers might like to use these three alternative models to see whether they give the same overall champion.

3. Conclusions

The examples above indicate just a few of the many ways in which mathematics can be used in sport. In (i) and (iii) mathematical analysis is used to provide a method of handicapping results and so obtain an overall winner, and in (ii) the mathematical model is used to improve technique.

References
1. I.A.A.F./F.I.A.A., *Progressive World Record Lists 1913–1977* (I.A.A.F. Publications, London, 1978).
2. J. B. Keller, Optimal velocity in a race, *Amer. Math. Monthly* **81** (1974), 474–480.
3. M. J. O'Carroll, On the relation between strength and bodyweight, *Res. Physical Educ.* **1**, No. 3 (1967), 6–11.

Stamps and Coins: Two Partition Problems

C. M. SHIU
Shell Centre for Mathematical Education
P. SHIU
Loughborough University of Technology

Christine and Peter Shiu are graduates of University College London. Christine started her career as a school teacher and she later joined Nottingham University as a research worker in mathematical education. She is currently employed on a Shell Centre project investigating the acquisition of general mathematical strategies by school children. Peter, who has twice contributed to *Mathematical Spectrum*, is a Lecturer in Mathematics at Loughborough University of Technology. His research interest is in number theory.

1. Stamps and coins

Consider the following two problems.

Problem 1. *How many ways are there of spending* £2 *on stamps of values* 2p, 3p *and* 5p?

Problem 2. *How many ways are there of making up* £1 *with* $\frac{1}{2}$p, 1p, 2p, 5p, 10p *and* 50p *coins?*

With patience we can obtain the answer to Problem 1 by exhaustive enumeration, but the method becomes hopelessly tedious when applied to Problem 2.

Let us consider the equations

$$2x_1 + 3x_2 + 5x_3 = 200 \tag{1}$$

and

$$x_1 + 2x_2 + 4x_3 + 10x_4 + 20x_5 + 100x_6 = 200 \tag{2}$$

where x_r $(1 \leq r \leq 6)$ are non-negative integers. It is clear that the numbers of solutions to these two equations are precisely the answers to our two problems.

2. Solution to problem 1

We shall deal with equation (1) first. Let us denote by $f(n)$ the number of solutions to

$$2x_1 + 3x_2 + 5x_3 = n,$$

so that our required answer is $f(200)$. It is not difficult to see that $f(n)$ is just the coefficient of t^n in the expression

$$(1 + t^2 + t^4 + t^6 + \cdots)(1 + t^3 + t^6 + t^9 + \cdots)(1 + t^5 + t^{10} + t^{15} + \cdots)$$

and therefore, for $|t| < 1$, we have

$$\sum_{n=0}^{\infty} f(n)t^n = \frac{1}{(1 - t^2)(1 - t^3)(1 - t^5)}. \tag{3}$$

In order to find $f(n)$ we resolve the right-hand side of this equation into partial fractions so that we can expand the individual terms into power series in t and collect the coefficients for t^n. Let us put

$$\omega = e^{2\pi i/3} \qquad \text{and} \qquad \theta = e^{2\pi i/5}$$

so that we have the factorisations

$$\begin{aligned}
1 - t^2 &= (1 - t)(1 + t)\\
1 - t^3 &= (1 - t)(1 - \omega t)(1 - \omega^2 t)\\
1 - t^5 &= (1 - t)(1 - \theta t)(1 - \theta^2 t)(1 - \theta^3 t)(1 - \theta^4 t).
\end{aligned}$$

We now write

$$\frac{1}{(1 - t^2)(1 - t^3)(1 - t^5)} = \sum_{m=1}^{3} \frac{A_m}{(1 - t)^m} + \frac{B}{1 + t} + \sum_{r=1}^{2} \frac{C_r}{1 - \omega^r t} + \sum_{r=1}^{4} \frac{D_r}{1 - \theta^r t}, \tag{4}$$

where A_m, B, C_r, D_r are constants. As we shall see in a moment, we only require the explicit knowledge of A_3 and A_2 for the purpose of finding $f(n)$. We have

$$\begin{aligned}
A_3 &= \lim_{t\to 1} \frac{(1 - t)^3}{(1 - t^2)(1 - t^3)(1 - t^5)}\\
&= \lim_{t\to 1} \frac{1}{(1 + t)(1 + t + t^2)(1 + t + t^2 + t^3 + t^4)} = \frac{1}{30}
\end{aligned}$$

and similarly

$$A_2 = -\lim_{t\to 1} \frac{d}{dt}\left\{\frac{(1 - t)^3}{(1 - t^2)(1 - t^3)(1 - t^5)}\right\} = \frac{7}{60}.$$

For completeness sake we also give

$$A_1 = 77/360, \qquad B = 1/8,$$
$$C_1 = -\omega^2/9, \qquad C_2 = -\omega/9,$$
$$D_1 = D_4 = \frac{(1-\theta)(1-\theta^4)}{25}, \qquad D_2 = D_3 = \frac{(1-\theta^2)(1-\theta^3)}{25}.$$

Now we have

$$(1-t)^{-1} = \sum_{n=0}^{\infty} t^n,$$

$$(1-t)^{-2} = \frac{d}{dt}(1-t)^{-1} = \sum_{n=0}^{\infty} (n+1)t^n,$$

$$(1-t)^{-3} = \frac{1}{2}\frac{d}{dt}(1-t)^{-2} = \sum_{n=0}^{\infty} \frac{1}{2}(n+1)(n+2)t^n,$$

$$(1+t)^{-1} = \sum_{n=0}^{\infty} (-1)^n t^n,$$

$$(1-\omega^r t)^{-1} = \sum_{n=0}^{\infty} \omega^{rn} t^n, \qquad (r = 1, 2)$$

and

$$(1-\theta^r t)^{-1} = \sum_{n=0}^{\infty} \theta^{rn} t^n, \qquad (r = 1, 2, 3, 4).$$

It therefore follows from (3) and (4) that

$$f(n) = \frac{A_3}{2}(n+1)(n+2) + A_2(n+1) + A_1 + (-1)^n B + \sum_{r=1}^{2} C_r \omega^{rn} + \sum_{r=1}^{4} D_r \theta^{rn}$$

and, with our calculations for A_3 and A_2, we now have that

$$\begin{aligned} f(n) &= \frac{(n+1)(n+9)}{60} + A_1 + (-1)^n B + \sum_{r=1}^{2} C_r \omega^{rn} + \sum_{r=1}^{4} D_r \theta^{rn} \\ &= \frac{n(n+10)}{60} + \left(\frac{3}{20} + A_1\right) + (-1)^n B + \sum_{r=1}^{2} C_r \omega^{rn} + \sum_{r=1}^{4} D_r \theta^{rn}; \end{aligned}$$

i.e.

$$f(n) = \frac{n(n+10)}{60} + \alpha_n, \tag{5}$$

where (α_n) is a bounded sequence. In fact α_n is made up of a constant term together with three terms which are periodic with periods 2, 3 and 5 so that (α_n) is periodic with period 30.†

† This means that $\alpha_{n+30} = \alpha_n$. (Note that 30 is the least common multiple of 2, 3, 5.)

From the direct enumeration of solutions we see that $f(20) = 11$ so that, from (5), $\alpha_{20} = 1$ and hence, by periodicity, $\alpha_{200} = 1$. Therefore, from (5)

$$f(200) = \frac{200 \times 210}{60} + 1 = 701.$$

3. An explicit formula for $f(n)$

By computing $f(n)$ for $n = 1, 2, \ldots, 30$ we can determine α_n for all n, and in fact we have

$$\begin{array}{rrrrrrrrrr} -11, & 36, & 21, & 4, & 45, & 24, & 1, & 21, & 9, & 40, \\ 9, & 21, & 1, & 24, & 45, & 4, & 21, & 21, & -11, & 60, \\ 9, & 16, & 21, & 24, & 25, & 24, & 21, & 16, & 9, & 60, \end{array}$$

for the first 30 values of $60\alpha_n$. We observe that $\alpha_n = -11/60$ if $n \equiv 1, 19 \pmod{30}$ and $0 < \alpha_n \leq 1$ otherwise. Let us denote by $[X]$ the integer part of X (that is, X is the largest integer which does not exceed X.) We now prove the following.

Lemma. Suppose that $X + \alpha$ is an integer. Then

$$X + \alpha = \begin{cases} [X] & \textit{if } -1 < \alpha \leq 0 \\ [X] + 1 & \textit{if } \ 0 < \alpha \leq 1. \end{cases}$$

Proof. Let us put $X = [X] + x$ so that $0 \leq x < 1$. We now have

$$X + \alpha = [X] + (x + \alpha)$$

where $x + \alpha$ must be an integer. Since $-1 < \alpha \leq 0$ implies $-1 < x + \alpha < 1$ and $0 < \alpha \leq 1$ implies $0 < x + \alpha < 2$, the required result follows.

It now follows from our lemma and the list of values for $60\alpha_n$ that, for all n,

$$f(n) = \left[\frac{n(n + 10)}{60}\right] + \beta_n$$

where

$$\beta_n = \begin{cases} 0 & \text{if } n \equiv 1, 19 \pmod{30} \\ 1 & \text{otherwise.} \end{cases}$$

4. Solution to Problem 2

We next deal with equation (2), and denote the number of its solutions by $f^*(200)$. It is convenient to reduce the number of variables in this problem from six to four by noting that $20x_5 + 100x_6 \leq 200$ implies $0 \leq x_5 \leq 10$ and $0 \leq x_6 \leq 2$ (the variables x_5 and x_6 correspond, of course, to the number of 10p and 50p coins used). Let us denote by $g(n)$ the number of solutions to the 'reduced' equation

$$x_1 + 2x_2 + 4x_3 + 10x_4 = n,$$

and consider Table 1.

TABLE 1

x_5	x_6	$200-(20x_5+100x_6)$	x_5	x_6	$200-(20x_5+100x_6)$
0	0	200	4	0	120
0	1	100	4	1	20
0	2	0	5	0	100
1	0	180	5	1	0
1	1	80	6	0	80
2	0	160	7	0	60
2	1	60	8	0	40
3	0	140	9	0	20
3	1	40	10	0	0

It is readily seen from this table that our desired solution to Problem 2 is given by

$$\begin{aligned} f^*(200) &= g(200) + 2g(100) + 3g(0) + g(180) + 2g(80) \\ &\quad + g(160) + 2g(60) + g(140) + 2g(40) \\ &\quad + g(120) + 2g(20) \\ &= 3g(0) + \sum_{m=1}^{5} \{2g(20m) + g(100+20m)\}. \end{aligned} \tag{6}$$

We now write $\phi = e^{2\pi i/10}$ so that we have the factorisations:

$$\begin{aligned} 1 - t^2 &= (1-t)(1+t), \\ 1 - t^4 &= (1-t)(1+t)(1-it)(1+it), \\ 1 - t^{10} &= (1-t)(1+t) \prod_{\substack{1\le r\le 9 \\ r\ne 5}} (1-\phi^r t); \end{aligned}$$

and we consider

$$\frac{1}{(1-t)(1-t^2)(1-t^4)(1-t^{10})} = \sum_{m=1}^{4} \frac{A_m}{(1-t)^m} + \sum_{m=1}^{3} \frac{B_m}{(1+t)^m} + \sum_{r=1,3} \frac{C_r}{1-i^r t} + \sum_{\substack{1\le r\le 9 \\ r\ne 5}} \frac{D_r}{1-\phi^r t}.$$

As before we need only calculate A_m and B_m with $m \geq 2$, and in fact we have

$$\begin{aligned} A_4 &= 1/80, \quad & A_3 &= 13/160, \quad & A_2 &= 39/160, \\ & & B_3 &= 1/160, \quad & B_2 &= 7/160. \end{aligned}$$

Since

$$(1-t)^{-4} = \sum_{n=0}^{\infty} \frac{1}{6}(n+1)(n+2)(n+3)t^n$$

we have, as in our analysis for $f(n)$, that

$$\begin{aligned} g(n) &= \frac{1}{6}A_4(n+1)(n+2)(n+3) \\ &\quad + \frac{1}{2}\{A_3 + (-1)^n B_3\}(n+1)(n+2) + \{A_2 + (-1)^n B_2\}(n+1) \\ &\quad + A_1 + (-1)^n B_1 + \sum_{r=1,3} C_r i^{rn} + \sum_{\substack{1 \le r \le 9 \\ r \ne 5}} D_r \phi^{rn} \\ &= X_n + \alpha'_n \end{aligned}$$

say, where

$$480X_n = \begin{cases} n^3 + 27n^2 + 212n & n \text{ even}, \\ n^3 + 24n^2 + 161n & n \text{ odd}, \end{cases}$$

and (α'_n) is periodic with period 20 (the least common multiple of 2, 4, 10).

By direct enumeration of solutions we find that $g(20) = 49$, and hence $\alpha'_{20} = 1$. Finally, then, we obtain

49, 242, 680, 1463, 2691, 4464, 6882, 10045, 14053, 19006

for $g(20m)$ with $m = 1, 2, \ldots, 10$, and from (6)

$$f^*(200) = 64703.$$

5. The partition problems

The two problems considered here are the so-called restricted partition problems. Thus, for instance, $f(n)$ is the number of ways of partitioning n into the restricted parts of 2, 3 and 5. The function

$$F(t) = \sum_{n=0}^{\infty} f(n)t^n$$

is called the *generating function* associated with this partition problem. We saw that restricted partition problems can be solved by resolving the generating function into partial fractions. However, the problem of unrestricted partitions, that is the determination of $p(n)$, the number of ways of writing n as a sum of positive integers with no restrictions on the parts, is very much more difficult. There is a recursion formula due to Euler for the computation of $p(n)$, but this is far too complicated to use when n is large, say $n \ge 1000$. With the definition $p(0) = 1$, the generating function for the unrestricted partition problem is

$$\sum_{n=0}^{\infty} p(n)t^n = \prod_{m=1}^{\infty} \frac{1}{1 - t^m};$$

but there is no easy method of retrieving $p(n)$ as an explicit function of n from this equation. In their famous collaboration in the early part of this century,

G. H. Hardy and the brilliant young Indian mathematician Ramanujan developed a powerful method of tackling this problem.† They proved many impressive results, for example,

$$\frac{\log p(n)}{\sqrt{n}} \to \pi \sqrt{\frac{2}{3}} \qquad \text{as } n \to \infty .$$

Some twenty years later H. Rademacher extended their work and determined $p(n)$ as the sum of a convergent series.

We point out that $p(n)$ is much larger than $f(n)$, $f^*(n)$ or $g(n)$. For example

$$p(200) = 3,972,999,029,388$$

a result obtained by MacMahon from Euler's recursion formula in order to give numerical verification of an even more accurate asymptotic formula of Hardy and Ramanujan than the one quoted above. The largest value of $p(n)$ yet computed is $p(14031)$, a number with 127 digits. This was done by D. H. Lehmer using Rademacher's formula in order to verify a conjecture of Ramanujan which includes the assertion that $p(14031)$ is divisible by $11^4 = 14641$.

† What J. E. Littlewood called 'the romantic story of the theorem' is reported in his fascinating review of the *Collected Papers of Srinivasa Ramanujan* in the *Mathematical Gazette* **14** (1929), 427–428.

Matters Mathematical (from *The Pirates of Penzance*; words by W. S. Gilbert).

I am the very pattern of a modern Major-Gineral,
I've information vegetable, animal, and mineral;
I know the kings of England, and I quote the fights historical,
From Marathon to Waterloo, in order categorical;
I'm very well acquainted, too, with matters mathematical,
I understand equations, both the simple and quadratical;
About binomial theorem I'm teeming with a lot o' news,
With interesting facts about the square of the hypotenuse.
I'm very good at integral and differential calculus,
I know the scientific names of beings animalculous.
In short, in matters vegetable, animal and mineral,
I am the very model of a modern Major-Gineral.

Dear Editor,

Desert island theorems

David Sharpe has chosen a fine time-intensive 'desert island theorem' in his choice of the formula

$$\frac{\pi^2}{6} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$$

(Volume 12 Number 3) if he aims to *calculate* π by this bit of Euler. If we use a primitive calculator (with a little application, which hurts no one), then 100 terms of the series give the value of π to be 3·132073. If we take the value of π to be 3·14159, this gives a deficiency of 0·00952. To 110 terms, we get 3·132934, so that the extra 10 terms decrease the deficiency by 0·000864. I have read somewhere that 100000 terms are necessary to produce π to 5 decimal places from this formula. Two matters are of interest to a non-mathematician such as myself: (1) How can this figure of 100000 be estimated? (2) What is the principle behind a computer program which makes the calculation?

You may wonder why an absolute non-mathematician like me can have an interest in *Spectrum*. There is usually a paragraph or two which pierces the gloom. This time I can share David Sharpe's astonishment at Euler's nerve in producing his 'proof'.

Yours sincerely,

T. C. Bagshaw
(Flat 2, Broadmead, Temple Street, Llandrindod Wells, Powys, LD1 5HW)

Dear Editor,

Pythagorean numbers

In his letter published in Volume 12 Number 3, Paul Heath shows ways of generating Pythagorean numbers from any positive whole number. It is also possible to obtain Pythagorean numbers from complex numbers. The square of any complex number with integer real and imaginary parts will yield a right triangle with integer sides if one takes the real part and the imaginary parts to be the two legs of the triangle. For

$$(a + bi)^2 = a^2 - b^2 + 2abi,$$

so that the two legs are $a^2 - b^2$ and $2ab$, where a, b are positive whole numbers. If one makes the proviso that $a > b$, this ensures that the $a^2 - b^2$ leg is of positive length. The hypotenuse squared will then be

$$(a^2 - b^2)^2 + (2ab)^2 = (a^2 + b^2)^2,$$

so the hypotenuse equals $a^2 + b^2$. If we insert, for instance, $a = 2$ and $b = 1$, we obtain the well-known 3, 4, 5 Pythagorean triangle.

Yours sincerely,

Barry Christian
(Department of Chemistry, Sheffield City Polytechnic)

Dear Editor,

Pythagorean numbers

Paul Heath, in a letter appearing on page 93 of Volume 12, Number 3, quotes two sets of Pythagorean triples. Expressed in terms of the positive integer n, his numbers are

$$a = n(n+2), \qquad b = \tfrac{1}{2}n(n^2+4n+3), \qquad c = \tfrac{1}{2}n(n^2+4n+5),$$
$$d = n^2+3n+1, \qquad e = \tfrac{1}{2}(n^4+6n^3+11n^2+6n), \qquad f = \tfrac{1}{2}(n^4+6n^3+11n^2+6n+2).$$

These identically satisfy $a^2 + b^2 = c^2$ and $d^2 + e^2 = f^2$. It is easily verified that the first set is

$$n[m, \tfrac{1}{2}(m^2-1), \tfrac{1}{2}(m^2+1)], \qquad \text{with } m = n+2,$$

whereas the second set is

$$[m, \tfrac{1}{2}(m^2-1), \tfrac{1}{2}(m^2+1)] \qquad \text{with } m = n^2+3n+1.$$

(His second set thus provides only a subset of his first set, if factors common to all triples are removed.)

It is not difficult to show that any Pythagorean triples, with common factors removed, can be written in the form

$$[pq, \tfrac{1}{2}(p^2-q^2), \tfrac{1}{2}(p^2+q^2)],$$

where p and q are mutually prime odd numbers. Paul Heath's formulae will only generate triples with $q = 1$, and so will not generate all possible triples. As his e and f differ by 1 and his b and c differ by 1 or 2 after common factors are removed, his formulae could never generate [33, 56, 65], for example.

Yours sincerely,

Bob Hale

(Division of Computing and Mathematics, Deakin University, Victoria 3217, Australia)

Dear Editor,

Two simple paradoxes

From time immemorial, along with the development of mathematics, paradoxes have been observed. I have observed the two paradoxes given here, which may amuse and puzzle your readers.

We first 'prove' that

$$\cdots = -3 = -2 = -1 = 0 = 1 = 2 = 3 = \cdots.$$

First, $a^0 = 1$ for $a \neq 0$, so $1^0 = 1$. But $1^m = 1$ for every positive or negative integer m. Hence

$$\cdots = 1^{-3} = 1^{-2} = 1^{-1} = 1^0 = 1^1 = 1^2 = 1^3 = \cdots.$$

Since the bases are equal, the powers must also be equal!

We now 'show' that either $\pi = 0$ or $i = 0$ (where $\pi = 3{\cdot}14\ldots$ and $i = \sqrt{(-1)}$). We know that

$$e^{\pi i} = \cos\pi + i\sin\pi = -1,$$

so

$$\ln(-1) = \pi i,$$

where $\ln(-1)$ is the logarithm of -1 to the base e. Hence

$$2\ln(-1) = 2\pi i,$$

so

$$\ln(-1)^2 = 2\pi i,$$
$$\ln 1 = 2\pi i,$$
$$0 = 2\pi i,$$

from which either $\pi = 0$ or $i = 0$.

Yours sincerely,

H. S. Krishnaswamy
(Department of Mathematics, College of Basic Sciences and Humanities, University of Agricultural Sciences, GKVK, Bangalore-560065, India)

(Ed: Perhaps other readers could send in their favourite paradoxes.)

Problems and Solutions

Sixth formers and students are invited to submit solutions to some or all of the problems below: the most attractive solutions will be published in subsequent issues. When writing to the Editorial Office, please state your full name and home address and also the postal address of your school, college or university.

Problems

13.4. (Submitted by R. J. Webster, University of Sheffield) A positive integer is said to be *normal* if the digits occurring in its decimal representation are all different (e.g. 123 is normal, but 122 is abnormal). Determine the positive integer or integers n such that there are equally many normal and abnormal numbers smaller than n.

13.5. (Submitted by L. Mirsky, University of Sheffield) Let n be a natural number ≥ 3, and let α, β, γ be complex numbers such that $\alpha^n = \beta^n = \gamma^n = 1, \alpha + \beta + \gamma = 0$. Prove that n is a multiple of 3.

13.6. Prove that $p(n) \geq 2^{[\sqrt{n}]}$ for all integers $n > 2$, where $[\sqrt{n}]$ denotes the integral part of $\sqrt{n}$. (See the article by C. M. Shiu and P. Shiu in this issue.)

Solutions to Problems in Volume 12, Number 3

12.7. Let a, b be positive real numbers and let n be an integer with $n > 2$. Show that

$$(a^n + b^n)^2 < (a^2 + b^2)^n.$$

Solution
Certainly

$$\left(\frac{a}{b}\right)^{n-2} + \left(\frac{b}{a}\right)^{n-2} > 1,$$

so

$$\begin{aligned}
&a^{2n-2}b^2 + b^{2n-2}a^2 > a^n b^n \\
&\Rightarrow n(a^{2n-2}b^2 + b^{2n-2}a^2) > 2a^n b^n \\
&\Rightarrow a^{2n} + na^{2n-2}b^2 + \cdots + na^2b^{2n-2} + b^{2n} > a^{2n} + 2a^n b^n + b^{2n} \\
&\Rightarrow (a^2 + b^2)^n > (a^n + b^n)^2.
\end{aligned}$$

12.8. Let k be an integer with $k \geq 2$ and let $n_1, \ldots, n_r$ $(r \geq 2)$ be integers $\geq k$ whose sum is denoted by n. Show that

$$\binom{n_1}{k} + \binom{n_2}{k} + \cdots + \binom{n_r}{k} < \binom{n - (r-1)(k-1)}{k},$$

where ( ) denotes the binomial coefficient.

Solution
We use induction on r. Consider the case $r = 2$. Put $n_1 = k + s_1, n_2 = k + s_2$, so that $s_1, s_2 \geq 0$. Then

$$\begin{aligned}
\binom{n-(k-1)}{k} &= \binom{s_1 + s_2 + k + 1}{k} \\
&= \frac{(s_1 + s_2 + k + 1)(s_1 + s_2 + k) \cdots (s_1 + s_2 + 2)}{k!} \\
&= \frac{(s_1 + 1)(s_1 + s_2 + k) \cdots (s_1 + s_2 + 2)}{k!} \\
&\quad + \frac{(s_2 + k)(s_1 + s_2 + k) \cdots (s_1 + s_2 + 2)}{k!} \\
&> \frac{(s_1 + 1)(s_1 + k)(s_1 + k - 1) \cdots (s_1 + 2)}{k!} \\
&\quad + \frac{(s_2 + k)(s_2 + k - 1) \cdots (s_2 + 1)}{k!} \\
&= \binom{n_1}{k} + \binom{n_2}{k},
\end{aligned}$$

which gives the result for $r = 2$. Now let $r \geq 2$, and assume the result for r. Then

$$\binom{n_1}{k} + \cdots + \binom{n_r}{k} + \binom{n_{r+1}}{k} < \binom{n - n_{r+1} - (r-1)(k-1)}{k} + \binom{n_{r+1}}{k}$$

$$< \binom{n - (r-1)(k-1) - (k-1)}{k}$$

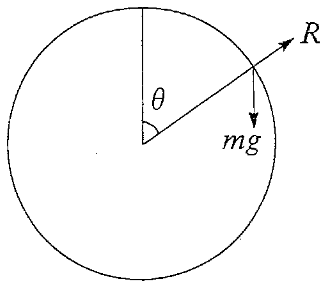
(using the result for the case $r = 2$)

$$= \binom{n - r(k-1)}{k},$$

which is the result for $r + 1$. The result follows for all integers $r \geq 2$, by induction.

12.9. A particle is slightly displaced from the top of a smooth, fixed sphere. Determine where the particle leaves the sphere.

Solution



Let the particle have mass m, and consider the particle when it has moved through an angle θ from the vertical, before it leaves the sphere. Let the reaction of the sphere on the particle be R. Then, if we resolve tangentially and radially, we have

$$mg \sin\theta = ma\ddot{\theta}, \qquad mg\cos\theta - R = ma\dot{\theta}^2,$$

where a is the radius of the sphere. Thus

$$g\sin\theta = a\ddot{\theta},$$
$$g(\sin\theta)\dot{\theta} = a\dot{\theta}\ddot{\theta},$$
$$-g\cos\theta = \tfrac{1}{2}a\dot{\theta}^2 + c,$$

where c is a constant. When $t = 0$ we have $\theta = 0$, $\dot{\theta} = 0$, and $c = -g$. Thus

$$a\dot{\theta}^2 = 2g(1 - \cos\theta).$$

The particle leaves the sphere when $R = 0$, i.e. when

$$g\cos\theta = a\dot{\theta}^2,$$

which gives

$$\cos\theta = 2(1 - \cos\theta),$$
$$3\cos\theta = 2$$
$$\cos\theta = \tfrac{2}{3}.$$

Thus the particle leaves the sphere when it has moved through an angle θ from the vertical, where $\cos\theta = \frac{2}{3}$.

Reason by Numbers. By PETER G. MOORE. Pelican Library of Business and Management, Penguin Books, Harmondsworth 1980. Pp. ii + 367 + xxvii. £1·95 (paperback).

I found this book well worth reading. The author is Professor of Statistics and Operational Research at the London Business School, and wrote the book because he found that many participants in the management courses there wanted to develop their quantitative skills, but were inhibited by the mathematics. However, he feels that it is the strict logical approach demanded by mathematical processes that in reality produces the blockage, and that greater efforts need to be placed on developing the skills of reasoning *per se*, and less on the teaching of mathematical theorems and techniques. In fact the bulk of the book does develop mathematical techniques, but always with a view to their sensible use to solve real problems.

The first chapter presents some 'limbering up' puzzles that simply require careful reasoning and the elimination of all but one possibility, and then gives the layout of the rest of the book. Chapter 2, Logical Overtures, is a brief introduction to logic, as far as the use of Venn diagrams to handle classical syllogisms. I felt this was the only disappointing part of the book, mainly because of the skimpiness of the treatment. Terms like 'major premise' and 'minor premise' appear without explanation (and are the wrong way round on page 33), the example about monopolies on page 30 is really an exercise in economics rather than logic, and altogether the chapter gives the impression, unlike any of the rest of the book, of having been written by someone with little real acquaintance with the subject. No use of it is made later on, and it could have been omitted, except that the Venn diagram technique is quite well explained and is useful to know.

Chapter 3 is on algorithms (and flowcharts), Chapter 4 deals with presenting and reading numerical information, and Chapter 5, 'Modern Metrology', includes the uses and abuses of indices like the Retail Price Index, and the problems involved in trying to give realistic definitions of productivity in various contexts. Welcome mention is also made of the way in which data errors can affect the result of a calculation, e.g. by changing the sign of the balance of trade!

The next two chapters apply algebra to business problems, including the discounting of cash flows (for which tables are provided) and linear programming. The latter is well explained, and I particularly liked the section on relaxing the constraints—what to do if more materials or more man-hours of labour become available, so as to make the best use of them.

Chapter 8 on statistical interpretation contains some good examples of the sort of pitfalls that are around when one draws conclusions too quickly from statistical information, and from then on the theme becomes probability and statistical inference. Particularly important is the notion of expectation, or the expected monetary value (EMV) of an action. To illustrate this, suppose that an oil 'wildcatter' has the chance of drilling a site at a cost of £100000, and that he estimates that he has a 30% chance of finding oil, and if so will gain £1000000, otherwise nothing. He has to decide whether or not to drill. If he doesn't, his gain is 0. If he does, his gain could be −£100000 with probability 0·7, or £900000 with probability 0·3. The EMV is an average of these weighted with their probabilities, namely

$$-\pounds100000 \times 0{\cdot}7 + \pounds900000 \times 0{\cdot}3 = \pounds200000.$$

This is the average gain he could expect if he took this action many times in similar circumstances. Of the two courses of action, drilling has the higher EMV; hence this would be the action to take provided the wildcatter (like a large company) could afford the loss of £100000; otherwise (like a small company) he would be wiser to use the 'minimax' criterion, i.e. act so as to minimise the maximum possible loss, by not drilling. This exemplifies an important theme in the book, that of making decisions in the face of uncertainty. There are also chapters on 'learning from samples' and 'sample surveys', which include a good deal of the mathematics used in drawing inferences about a population from information about a sample, and one on simulation, a widely used and powerful technique.

The final chapter returns to decision-making and allows for the possibility of obtaining extra information about the uncertainties. Thus suppose the wildcatter above could commission a survey that would tell him with certainty whether or not he would strike oil if he drilled. How much would it be worth paying for this? The survey has a 30 % chance of giving a positive answer, in which case a gain of £900000 is sure; otherwise the gain is 0. The EMV is now £900000 × 0·3 + £270000, which is £70000 more than before. Thus the expected value of perfect information in this situation is £70000 (surprisingly low!) and on an EMV basis this is the most that is worth paying for the survey.

In practice a survey is not likely to yield perfect information: it might have an 80 % chance, say, of giving a positive indication when a well drilled on the site would be 'wet' (yield oil) and a 10 % chance of giving the same when in fact a well would be 'dry'. In that case one has to work out what probability the well has of being wet if the survey gives a positive result, and to do this Bayes' theorem is introduced. The proof is omitted, but is virtually derived in the author's solution of his wildcatter's problem (using different figures from mine above) on pages 361–2. By this time the reader is ready to tackle some quite meaty problems in decision-making under uncertainty.

I found the book well written and very readable, though with plenty of content that demands careful attention. I particularly liked the way that the problems and their solutions were related to reality: it was obvious that the author has had experience of the world of business. (He is a director of several companies and consultant to others.) This added both interest and relevance to the work. The mathematics is sound and well explained, and so is most of the statistics, though inevitably some formulae, like those for the binomial and Poisson distributions, have to be given without derivation.

On page 251 the author proposes to calculate how large a sample is needed to determine whether a population proportion has changed 'significantly' from 25 %, 'significantly' here meaning by at least 5 %. This use of the word is radically different from that made in this context by statisticians, for whom it refers not to the size of a change in the population proportion but to the discrepancy between a supposed population proportion and an observed sample proportion. Also the calculation that follows is achieving something quite different from the objective stated. The sample size of 310 allows a confidence interval of the form $P \pm 5\%$ to be constructed, with 95 % chance that it will contain the new population proportion, but if for instance P comes to 23 % then the confidence limits are 18 % and 28 %, and we still don't know whether a 'significant' change from 25 % has occurred. In fact no sample size can *guarantee* to detect whether a proportion has changed by as much as 5 %. The author's non-standard use of 'significant' occurs again at the bottom of page 261, in a section on 'testing for significance' which is otherwise clear and helpful.

A non-mathematical student reading this book on his own will struggle in some places; the assurance in the preface that 'algebra of the O or CSE level is only introduced in a few places' is somewhat optimistic; but this is probably inevitable if one wants to cover as much ground as this book does. Some of the answers (at the end of the book) to verbal questions are unsatisfyingly brief, and I disagreed with those to Nos. 6.4 and 9.9. But these are minor complaints; in general the exercises and their answers are extremely useful. There are several tables, and even nomograms for finding standard errors without calculation. Such misprints as I noticed were mostly obvious enough not to cause much trouble, though there is a mistake in Example 6.5, page 140, which may perplex the less confident reader.

I must apologise for the length of this review, but feel that the book warrants it, and that it can be strongly recommended to anyone who wants to acquaint himself with its subject-matter.

Royal Grammar School, High Wycombe C. W. PURITZ

Catastrophe Theory. By ALEXANDER WOODCOCK and MONTE DAVIS. Penguin Books, Harmondsworth. Pp. 171. Paperback £1·25.

Catastrophe theory is part of pure mathematics concerning the behaviour of maxima and minima of potential functions. Its name derives from the fact that a continuous change in the potential function can induce a discontinuous, or 'catastrophic', change in the relative position of the maxima or minima.

Many phenomena of natural science and social science are determined in exactly this way, so that the pure mathematics of catastrophe theory provides a possible model for making sense of the discontinuous changes which these phenomena display. If the potential function is known, as in many problems in physics and engineering, then catastrophe theory can be used to predict the resultant behaviour. If the potential function is unknown, as in many problems in psychology and biology, then the model can be used to guess possible potential functions from the observed behaviour.

This book is a British reprint of a non-mathematical description of catastrophe theory written in the U.S.A. at the end of 1977 during a controversy about the use of catastrophe theory which was largely generated by petty jealousies and ludicrous misunderstandings. Too much of the book is devoted to the controversy and too little to the actual problem of fitting the pure mathematics to observed data. Too much describes rather vague 'applications' to biology, psychology, sociology and economics, whereas too little space is given to the striking successes of catastrophe theory in physics and engineering.

The authors are too anxious to be even-handed in their praise and criticism; as a result they state unfair criticisms of Christopher Zeeman which could be applied more tellingly to their own book. They should write a more objective and up-to-date second edition.

University of Warwick — R. L. E. SCHWARZENBERGER

A Structured Introduction to Numerical Mathematics. By P. J. HARTLEY and A. WYNN-EVANS. Stanley Thornes (Publishers) Ltd., Cheltenham. Pp. viii + 456. Soft covers £6·75.

As someone who likes to search through a variety of sources in order to develop a more personal approach to the teaching of a course, I have an inherent dislike of 'structured' books containing 'lessons'. Too frequently the lesson is rather formal, does not contain exactly what I want, and spends too much time developing techniques, without provoking any real thinking in the student.

I was therefore pleasantly surprised to find that this book poses many a question for the student as the theory is developed through worked examples. Exercises are plentiful and the answers often include some explanation. Further reading is suggested at the end of each lesson, with a brief indication of the level and approach.

It is a pity that there is little use made of the computer. Print-outs of computer runs of BASIC programs, for example, would have enlightened the student even if his knowledge of programming was limited. However, the book covers the various techniques well, apart from this one deficiency.

The book covers five major areas: iterative methods, solution of simultaneous equations, interpolation and approximation, integration and ordinary differential equations. The technique of least squares is not included.

Most topics are introduced intuitively with geometrical interpretations used where possible. This approach would suit most students on an initial course about numerical methods at first-year undergraduate level. An analytical consideration is injected at some point during each section, though I feel this is done rather inconsistently for a structured approach. For example, though the sections on iteration and integration are introduced geometrically, before any attempt to consider the topics analytically, the section on interpolation starts with a large section concerning Taylor's theorem and the remainder term, which could be off-putting.

In general the book seems good value for money, with 456 pages, and it is a book which students can *use*, in order to familiarise themselves with the numerical techniques and gain some insight into the theoretical background.

St Mary's College, Strawberry Hill K. M. GILBERT

Applied Statistics with Probability. By J. S. MILTON and J. J. CORBET. D. Van Nostrand Company, Wokingham. Pp. xi + 487. Hard Covers £11·95.

There are many statistics books being published in the United States with essentially similar contents. Their distinguishing features are that they deal with discrete probability distributions, spend some time on permutations and combinations linked with probability, and emphasise estimation and significance testing of means of samples using normal distribution theory. They do not use calculus, and hence the work on continuous distributions is weak. This book goes further than most, in that it also includes tests and estimation on sample variances, including the use of the F test, some simple analysis of variance, the use of linear contrasts and some elementary non-parametric tests. There is no reference to the Poisson distribution.

The text is slow-moving, and there are a large number of examples with selected answers at the back, so the book could be used for self-instruction. Two idiosyncracies were noted: the use of T instead of the more usual t for Student's distribution, and the use of $\cap$ for 'has distribution' as in '$Z \cap N(0, 1)$', meaning 'Z has normal distribution with mean 0 and variance 1'.

In the preface the authors state that the student needs to have examples and exercises that present the same concept in different settings. If interpreted literally, the need is met in the text, in that the verbiage round the calculation does refer to many different areas of application. On the other hand, a closer examination of the data reveals that most have been invented for the purpose of the calculation. This raises an important matter of principle; which is the more important, the technique or the practical implications of the data? The use of invented data leads inexorably to the conclusion that it is the techniques that are more important, since no meaning can be given to the results. The use of real data, although much more difficult to obtain, would have shown that the techniques are important in helping to draw real inferences in practical situations. The conclusions would then have real meaning and give extra motivation for learning the techniques.

University of Sheffield P. HOLMES

Contents

ISSN 0025-5653

PRICES (*postage included*)

Prices for Volume 13 (Issues Nos. 1, 2, and 3):

Subscribers in Britain and Europe: £2·00
Subscribers overseas: £4·00 (US$9.00; $A8.00)
(These prices apply even if the order is placed by an agent in Britain.)
A discount of 10% is allowed on all orders for five or more copies.

Back issues:

All back issues are available; information concerning prices may be obtained from the Editor.

Enquiries about rates, subscriptions and advertisements should be directed to:

Editor—*Mathematical Spectrum*,
Hicks Building,
The University,
Sheffield S3 7RH, England.

Printed in England by Galliard (Printers) Ltd, Great Yarmouth