Probability problems for Sigma - part 1

Andy Mitchell

October 18, 2021

Hello Sigma members!
After a long gap I have, at last, managed to find some time to revisit some mathematics and started reading an old book I have on probability. That book is “Probability - an Introduction” by Grimmet and Welsh (1985) from which I’ve investigated some of the problems/exercises from the first chapter which I hope you’ll find interesting.
I’m starting this submission with a quick guide on recurrence relations (also based on a similar section from the same book) which seem to be useful in many situations when looking at probabilty problems. Hopefully, I’ll be looking at some extra problems for a subsequent newsletter.

1 A quick guide to solving linear recurrence relations

1.1 Introduction

This section only deals with techniques for solving recurrence relations (or “difference equations”) where we have a sequence of numbers $x_0, x_1, \dots$ satisfying the relationship

$$a_0 x_{n+k} + a_1 x_{n+k-1} + \cdots + a_k x_n = g(n) \qquad \text{for } n = 0, 1, 2 \dots \tag{R1}$$

where $a_0, a_1, \dots a_k$ are given real numbers and both $a_0 \neq 0$ and $a_k \neq 0$, and $g(n)$ is some polynomial function of n (or a constant).
To get an explicit solution, we also assume that we know the initial values $x_0, x_1, \dots x_{k-1}$ (although any k known values would suffice).
The first step is to find the roots of the *auxiliary equation* (sometimes called the ’characteristic equation’) :

$$a_0 \theta^k + a_1 \theta^{k-1} + \cdots + a_{k-1}\theta + a_k = 0 \tag{1}$$

If all of the roots (including complex roots) $\theta_1, \theta_2, \dots \theta_k$ are distinct, then the general solution of (R1) (where $g(n) \equiv 0$ for now) is :

$$x_n = c_1 \theta_1^n + c_2 \theta_2^n + \cdots + c_k \theta_k^n \qquad \text{for } n = 0, 1, 2 \dots \tag{2}$$

and $c_1, c_2, \dots c_k$ are arbitrary constants which can be found by equating the first k values with the known initial values, and solving the simultaneous equations.
If any roots are the same (and occur m times), then the term for that root θ will instead be of the form $(s_1 + s_2 n + \cdots + s_m n^{m-1})\theta^n$
Finally, for $g(n) \not\equiv 0$ we firstly solve as above for $g(n) \equiv 0$ to obtain the *complementary solution* $x_n = \kappa_n$, then by any devious/clever means required, we find a *particular solution* for our particular $g(n)$, $x_n = \pi_n$
The general solution is then simply $x_n = \kappa_n + \pi_n$.

1.2 Example 1

Solve $x_{n+2} - 7x_{n+1} + 10x_n = 0$ where $x_0 = 2$ and $x_1 = 3$.
The auxiliary equation is $\theta^2 - 7\theta + 10 = 0 \implies \theta_1 = 2, \theta_2 = 5$ (the root order is not important).

The general solution is therefore of the form :

$$x_n = c_1 2^n + c_2 5^n \tag{3}$$

Setting $n = 0$ and $n = 1$ respectively, we have :

$$\begin{aligned} 2 &= c_1 + c_2 \\ 3 &= 2c_1 + 5c_2 \end{aligned}$$

which leads to $c_1 = \frac{7}{3}$ and $c_2 = \frac{-1}{3}$
The general solution is therefore :

$$x_n = \frac{7}{3} 2^n - \frac{1}{3} 5^n \tag{4}$$

1.3 Example 2

Solve $x_{n+3} - 5x_{n+2} + 8x_{n+1} - 4x_n = 0$ where $x_0 = 0$, $x_1 = 3$, $x_2 = 13$
The auxiliary equation is $\theta^3 - 5\theta^2 + 8\theta - 4 = 0 \implies \theta_1 = 2, \theta_2 = 2, \theta_3 = 2$ (again the root order is not important).
The general solution is therefore of the form :

$$x_n = c_1 1^n + (c_2 + c_3 n) 2^n \tag{5}$$

The boundary conditions can then be used to determine $c_1 = 1$, $c_2 = -1$ and $c_3 = 2$

1.4 Example 3

Solve $x_{n+2} - 5x_{n+1} + 6x_n = 4n + 2$ where $x_0 = 5$, $x_4 = -37$
As above, the complementary solution is of the form

$$x_n = c_1 2^n + c_2 3^n \tag{6}$$

By guesswork, trial and error, magic, instinct, or luck, we can find that $x_n = 2n + 4$ for $n = 0, 1, 2, \ldots$ is a particular solution, so the general solution is therefore

$$x_n = c_1 2^n + c_2 3^n + 2n + 4 \tag{7}$$

The boundary conditions can then be used to determine $c_1 = 2$ and $c_2 = -1$.

2 Grimmet and Welsh Chapter 1 Problem 1

2.1 Question

A fair die is thrown n times. Show that the probability that there are an even number of sixes is

$$\mathbf{P}_n := \frac{1}{2}\left[1+\left(\frac{2}{3}\right)^n\right]$$

2.2 Some general observations

Looking at specific values of n seems a good idea to get a feel for the behaviour of $\mathbf{P}_n$ and to validate any formulae we obtain.
$\mathbf{P}_0 = 0$, $\mathbf{P}_1 = \frac{5}{6}$, $\mathbf{P}_2 = \frac{1}{36} + \frac{25}{36} = \frac{26}{36}$
For very large n, the number of sixes will be broadly spread around the value $\frac{n}{6}$ with broadly similar amounts of even and odd values, so we should also expect

$$\lim_{n\to\infty} \mathbf{P}_n = \frac{1}{2}$$

2.3 Answer 1

A direct approach!
We will use this standard result : If $A_1, A_2, \ldots$ are disjoint, i.e. $A_i \cup A_j = \emptyset$ for $i \neq j$:

$$\mathbf{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbf{P}(A_i) \tag{F1.1}$$

Firstly, consider when n is even (only to save messy-looking equations) :
Let ${}_i\mathbf{P}_n$ be the probability of throwing i sixes from n throws.
Using (F1.1) :

$$\mathbf{P}_n = {}_0\mathbf{P}_n + {}_2\mathbf{P}_n + \ldots + {}_n\mathbf{P}_n$$

But ${}_i\mathbf{P}_n = \binom{n}{i}\left(\frac{1}{6}\right)^i\left(\frac{5}{6}\right)^{n-i}$, so

$$\mathbf{P}_n = \binom{n}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^n + \binom{n}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{n-2} + \ldots + \binom{n}{n}\left(\frac{1}{6}\right)^n\left(\frac{5}{6}\right)^0$$

To evaluate this, note that :

$$(1+x)^n + (1-x)^n = 2\left(\binom{n}{0}x^0 + \binom{n}{2}x^2 + \ldots + \binom{n}{n}x^n\right)$$

Setting $x = \frac{1}{5}$:

$$\frac{1}{2}\left[\left(\frac{6}{5}\right)^n + \left(\frac{4}{5}\right)^n\right] = \binom{n}{0}\left(\frac{1}{5}\right)^0 + \binom{n}{2}\left(\frac{1}{5}\right)^2 + \ldots + \binom{n}{n}\left(\frac{1}{5}\right)^n$$

Multiplying by $\left(\frac{5}{6}\right)^n$ we get :

$$\mathbf{P}_n = \left(\frac{5}{6}\right)^n \times \left(\frac{1}{2}\right)\left(\left(\frac{6}{5}\right)^n + \left(\frac{4}{5}\right)\right)^n = \frac{1}{2}\left[1+\left(\frac{2}{3}\right)^n\right]$$

For when n is odd, very similar logic applies. The revised equations corresponding to above are :

$$\mathbf{P}_n = {}_0\mathbf{P}_n + {}_2\mathbf{P}_n + \ldots + {}_{n-1}\mathbf{P}_n$$

$$\mathbf{P}_n = \binom{n}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^n + \binom{n}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{n-2} + \ldots + \binom{n}{n-1}\left(\frac{1}{6}\right)^{n-1}\left(\frac{5}{6}\right)^1$$

$$(1+x)^n + (1-x)^n = 2\left(\binom{n}{0}x^0 + \binom{n}{2}x^2 + \ldots + \binom{n}{n-1}x^{n-1}\right)$$

and the same result is obtained.

2.4 Answer 2

Using a recurrence relation.
If we throw a die $n+1$ times then to obtain an even number of sixes, either the first throw is not a six (with probabilty $5/6$) and we need an even number of sixes from the remaining n throws, or the first throw is a six (with probability $1/6$) and we need an odd number of sixes from the remaining n throws.
So, conditioning on the first throw :

$$\mathbf{P}_{n+1} = \frac{5}{6}\mathbf{P}_n + \frac{1}{6}(1-\mathbf{P}_n) \Rightarrow \mathbf{P}_{n+1} = \frac{1}{6} + \frac{2}{3}\mathbf{P}_n$$

Using "standard" methods, the auxiliary equation is $\theta - \frac{2}{3} = 0$, yielding the complementary solution :

$$A\left(\frac{2}{3}\right)^n$$

and a particular solution $\mathbf{P}_n = \frac{1}{2}$
So, the general solution is $\mathbf{P}_n = \frac{1}{2} + A\left(\frac{2}{3}\right)^n$, but as we know $P_0 = 1$, then $A = \frac{1}{2}$.
Therefore, as before :

$$\mathbf{P}_n = \frac{1}{2}\left[1 + \left(\frac{2}{3}\right)^n\right]$$

2.5 Answer 3

Given that we are provided with the answer, we can use induction on n.
This is easiest but least pleasing (as induction often is) as we don't really gain any insight into why it is true....
The statement is clearly true for $n = 0$ as $\mathbf{P}_0 = 0$.
Assume it is true for $n = k$. Therefore :

$$\mathbf{P}_k = \frac{1}{2}\left[1 + \left(\frac{2}{3}\right)^k\right]$$

But, as in Answer 2,

$$\mathbf{P}_{k+1} = \frac{1}{6} + \frac{2}{3}\mathbf{P}_k = \frac{1}{6} + \frac{2}{3} \times \frac{1}{2} \times \left[1 + \left(\frac{2}{3}\right)^k\right] = \frac{1}{2}\left[1 + \left(\frac{2}{3}\right)^{k+1}\right]$$

so it is also true for $n = k+1$ and so true for all $n \geq 0$.

2.6 Final comments

As expected, the formula does indeed tend to $\frac{1}{2}$ as $n \to \infty$. Also, an obvious consequence is that the probability that there are an odd number of sixes in n throws is

$$\mathbf{P}_n = \frac{1}{2}\left[1 - \left(\frac{2}{3}\right)^n\right]$$

3 Grimmet and Welsh Chapter 1 Exercise 27

3.1 Question

A biased coin shows heads with probability $p = 1 - q$ whenever it is tossed.
Let u_n be the probability that, in n tosses, no pair of heads occur successively. Show that for $n \geq 1$

$$u_{n+2} = qu_{n+1} + pqu_n$$

and find u_n by the usual method if $p = \frac{2}{3}$

3.2 Answer 1

We will be using the "Partition Theorem" :
If $B_i \cap B_J = \emptyset$ and $\cup_i(B_i) = \Omega$ and $\mathbf{P}(B_i) > 0 \qquad \forall i$:

$$\mathbf{P}(A) = \sum_i \mathbf{P}(A \mid B_i)\mathbf{P}(B_i) \qquad \text{for all } A \tag{F1.9}$$

Let A_n represent the event that, in n tosses, no pair of heads occur successively Let B_i be the event that the first $(i-1)$ tosses are heads ("H") and the ith toss is tails ("T") (and $i \geq 1$)
So B_1 represents the events $T\ldots$,
B_2 represents the events $HT\ldots$,
B_3 represents the events $HHT\ldots$ etc.
Note that if Ω represents all possible events

$$\bigcup_i B_i = \Omega \text{ and } B_i \cap B_j = \emptyset \text{ for all } i \neq j$$

So, we can use the Partition Theorem (F1.9) as follows.

$$u_{n+2} = \mathbf{P}(A_{n+2}) = \mathbf{P}(A_{n+2} \mid B_1)\mathbf{P}(B_1) + \mathbf{P}(A_{n+2} \mid B_2)\mathbf{P}(B_2) + \ldots$$

But, $\mathbf{P}(A_{n+2} \mid B_1)$ looks at the event given a beginning "T", which 'resets' the starting state to having $n+1$ tosses in which to get no pair of heads (since the tail does not affect the situation for subsequent throws). This also applies to $\mathbf{P}(A_{n+2} \mid B_2)$.
$\mathbf{P}(A_{n+2} \mid B_3)$, $\mathbf{P}(A_{n+2} \mid B_4)\ldots$ are all 0 since the outcome of no pair of heads has failed for all of them. So,

$$= u_{n+1}q + u_n pq + 0 + 0 + \ldots$$

For the second part, we substitute $p = \frac{2}{3}$ to obtain

$$u_{n+2} = \frac{1}{3}u_{n+1} + \frac{2}{9}u_n \implies 9u_{n+2} - 3u_{n+1} - 2u_n = 0$$

and we also note that $u_1 = 1$, and $u_2 = \frac{5}{9}$
Using "standard" methods, the auxiliary equation is $9\theta^2 - 3\theta - 2 = 0 \to (3\theta+1)(3\theta-2) = 0$, yielding the general solution :

$$u_n = A\left(\frac{-1}{3}\right)^n + B\left(\frac{2}{3}\right)^n \text{ for some } A \text{ and } B.$$

$$u_1 = 1 \implies -\frac{1}{3}A + \frac{2}{3}B = 1 \to 2B - A = 3$$

$$u_2 = \frac{5}{9} \implies A + 4B = 5$$

which leads to

$$B = \frac{4}{3} \text{ and } A = -\frac{1}{3}$$

$$\therefore u_n = \frac{4}{3}\left(\frac{2}{3}\right)^n - \frac{1}{3}\left(\frac{-1}{3}\right)^n \text{ for } n \geq 1 \tag{8}$$

There are other approaches considered later.

3.3 Comments

Below is a table of probabilities for the first few values of n. As it is easy to make mistakes deriving formulae, it's a good idea to verify results by using a separate mechanism - in this case, a simple Java program which 'makes', for each n, n tosses of a coin millions of times and counts the proportion which don't contain 2 consecutive heads.
As you can see, the results suggest that the answers above are (probably) correct!
Also note the successvie ratio of probabilities for increasing values of n which, as expected, converge to the dominant power 2/3 in (8) above.

3.4 Alternative solution

We can also get the same result by letting v_n be the probability that, in n tosses, a pair of heads does occur successively. Defining A_n^* as the event that, in n tosses, a pair of heads does occur successively, and B_i exactly as before,

$$\begin{aligned}
v_{n+2} &= \mathbf{P}(A_{n+2}^*) = \mathbf{P}(A_{n+2}^* \mid B_1)\mathbf{P}(B_1) + \mathbf{P}(A_{n+2}^* \mid B_2)\mathbf{P}(B_2) + \dots \\
&= v_{n+1}q + v_n pq + 1 \cdot pq + 1 \cdot p^2 q + 1 \cdot p^3 q + \dots \\
&= v_{n+1}q + v_n pq + \frac{p^2 q}{1-p} \quad = v_{n+1}q + v_n pq + p^2
\end{aligned}$$

from which we can either substitute $u_n = 1 - v_n$ to get the previous relation, or proceed to solve it for a particular value of p and substitute later...
Solving for $p = \frac{2}{3}$ we obtain

$$v_n = 1 - \frac{4}{3}\left(\frac{2}{3}\right)^n + \frac{1}{3}\left(\frac{-1}{3}\right)^n \equiv 1 - u_n \tag{9}$$

4 Grimmet and Welsh Chapter 1 Problem 7

4.1 Question

Two people toss a fair coin n times. Show that the probability that they throw an equal number of heads is

$$\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}$$

4.2 Answer 1

The hints section of the book suggests this.
If X and Y are the numbers of heads tossed by each person :

$$\mathbf{P}(X = Y) = \sum_k \mathbf{P}(X = k)\mathbf{P}(Y = k) = \sum_k \mathbf{P}(X = k)\mathbf{P}(Y = n - k)$$

The last statement works because the coin is fair, so this is essentially saying that k heads has the same probability as k tails, which is the same event as $n - k$ heads. But,

$$\sum_k \mathbf{P}(X = k)\mathbf{P}(Y = n - k) = \mathbf{P}(X + Y = n)$$

which is the probability of, regardless of how they are distributed amongst each person, n heads result from $2n$ tosses, which is just $\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}$

n	$\frac{4}{3}\left(\frac{2}{3}\right)^n - \frac{1}{3}\left(\frac{-1}{3}\right)^n$	Ratio	Java program
1	1		1
2	0.5555555556	0.555556	0.55553827
3	0.4074074074	0.733333	0.407404815
4	0.2592592593	0.636364	0.25929586
5	0.1769547325	0.68254	0.176948485
6	0.1165980796	0.658915	0.116604615
7	0.0781893004	0.670588	0.078175775
8	0.0519737845	0.664717	0.05198466
9	0.0346999949	0.667644	0.034695565
10	0.0231163949	0.666179	0.023142995
11	0.0154165749	0.666911	0.01540314
12	0.0102758349	0.666545	0.01027086

4.3 Answer 2

Here is a more direct (and longer!) approach.

$$\mathbf{P}(X=Y)=\sum_k \mathbf{P}(X=k)\mathbf{P}(Y=k)=\sum_k \frac{1}{2^k}\frac{1}{2^{n-k}}\binom{n}{k}\frac{1}{2^k}\frac{1}{2^{n-k}}\binom{n}{k}=\frac{1}{2^{2n}}\sum_{k=0}^{n}\binom{n}{k}^2$$

Now,

$$(1+x)^n=\binom{n}{0}x^0+\binom{n}{1}x^1+\binom{n}{2}x^2+\cdots+\binom{n}{n-1}x^{n-1}+\binom{n}{n}x^n \text{ , and}$$

$$(x+1)^n=\binom{n}{0}x^n+\binom{n}{1}x^{n-1}+\binom{n}{2}x^{n-2}+\cdots+\binom{n}{n-1}x^1+\binom{n}{n}x^0$$

If we multiply these two equations, but only look at the coefficients of x^n we obtain :

$$\binom{2n}{n}x^n=x^n\left(\binom{n}{0}^2+\binom{n}{1}^2+\binom{n}{2}^2+\cdots+\binom{n}{n-1}^2+\binom{n}{n}^2\right)$$

and the result follows immediately.

4.4 Pascals triangle

The above equations have an interesting 'visual' appeal when we look at Pascal's triangle.

n=0									1								
n=1								1		1							
n=2							1		2		1						
n=3						1		3		3		1					
n=4					**1**		**4**		**6**		**4**		**1**				
n=5				1		**5**		**10**		**10**		**5**		1			
n=6			1		6		**15**		**20**		**15**		6		1		
n=7		1		7		21		**35**		**35**		21		7		1	
n=8	1		8		28		56		**70**		56		28		8		1

Looking at the inverted triangle of bold red values starting at the line $n=4$ and reaching the value 70 at the lowest point (where $n=8$), we now know that
$1^2+4^2+6^2+4^2+1^2=70$
This works for any such inverted triangle.

4.5 Large n

We would certainly expect the value of $\binom{2n}{n}x^n$ to approach 0 as $n\to\infty$, since the chance of the same number of heads for an enormous number of throws seems remote.
The value also seem to decrease "quite slowly", e.g. the values for $n=8$ and $n=32$ are around $\frac{1}{5}$ and $\frac{1}{10}$ respectively.
Using Stirling's formula :

$$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \implies \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\sim\left(\frac{\sqrt{2\pi 2n}(\frac{2n}{e})^{2n}}{\sqrt{2\pi n}(\frac{n}{e})^n\times\sqrt{2\pi n}(\frac{n}{e})^n}\right)\left(\frac{1}{2}\right)^{2n}$$

$$=\frac{1}{\sqrt{\pi n}}$$

This is a very good approximation for the actual values, even for small n.

4.6 What about an unfair coin?

We cannot use the sneaky trick from Answer 1, so we can perhaps attempt a direct approach again... We now use a biased coin which shows heads with probability $p = 1 - q$ whenever it is tossed.
Now,

$$\mathbf{P}(X = Y) = \sum_k \mathbf{P}(X = k)\mathbf{P}(Y = k) = \sum_k \left(\binom{n}{k}\right)^2 p^2 k q^2 n - k$$

If we try a similar trick to evaluate this, we would start with :

$$(1 + p^2 x)^n = \binom{n}{0} x^0 + \binom{n}{1} p^2 x^1 + \binom{n}{2} p^2 x^2 + \cdots + \binom{n}{n-1} p^{2n-2} x^{n-1} + \binom{n}{n} p^{2n} x^n \text{ , and}$$

$$(q^2 x + 1)^n = \binom{n}{0} q^{2n} x^n + \binom{n}{1} q^{2n-2} x^{n-1} + \binom{n}{2} q^{2n-4} x^{n-2} + \cdots + \binom{n}{n-1} q^2 x^1 + \binom{n}{n} x^0$$

and we could try to determine the coefficient of x^n in $(1 + p^2 x)^n (1 + q^2 x)^n$
Unfortunately, there appears to be no naturally simple answer to this (except for the sum we are trying to determine in the first place).
I suspect there is no closed formula for the probability for the unfair coin, but I'm very happy to be corrected by any readers!