Probability

Andy Mitchell

October 18, 2021

Contents

1 Formulae from "Probability - an Introduction" by Grimmet and Welsh (1985)

1.1 From Chapter 1 and a few extras

If $A_1, A_2, \ldots$ are disjoint, i.e. $A_i \cup A_j = \emptyset$ for $i \neq j$:

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

$A \setminus B$ is the *difference* of sets A and B i.e. all points of the sample space Ω which are in A but not in B.

$$A \setminus B := A \cap (\Omega \setminus B) \tag{F1.2}$$

$$\mathbf{P}(A) = \mathbf{P}(A \setminus B) + \mathbf{P}(A \cap B) \tag{F1.3}$$

For all sets of events A, B :

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \tag{F1.4}$$

And more generally :

$$\mathbf{P}(\bigcup_{1 \leq i \leq n} A_i) = \sum_{1 \leq i \leq n} \mathbf{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathbf{P}(A_i \cap A_j) + \ldots + (-1)^{n+1} \mathbf{P}(\bigcap_{1 \leq i \leq n} A_i) \tag{F1.5}$$

The *conditional* probability of A given B is (if $\mathbf{P}(B) > 0$) :

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \tag{F1.6}$$

Events A and B are *independent* if $\mathbf{P}(A \mid B) = \mathbf{P}(A)$ and $\mathbf{P}(B \mid A) = \mathbf{P}(B)$. From (F1.6), this implies :

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B) \tag{F1.7}$$

Events A_i (where $i \in I$) are *independent* if for ALL subsets $J \subseteq I$:

$$\mathbf{P}(\bigcap_{i \in J} A_i) = \prod_{i \in J} \mathbf{P}(A_i) \tag{F1.8}$$

1.2 Partition Theorem

If $B_i \cap B_J = \emptyset$ and $\cup_i(B_i) = \Omega$ and $\mathbf{P}(B_i) > 0 \qquad \forall i$:

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

1.3 Formulae involving p and q

This section includes some random identities involving the probabilities p and $q = 1 - p$:

$$p^2 + q^2 \equiv 1 - 2pq \tag{PQ1}$$

$$p + q^2 \equiv q + p^2 \tag{PQ2}$$

$$q + pq + p^2 \equiv 1 \tag{PQ3}$$

$$p^3 + q^3 \equiv 1 - 3pq \tag{PQ4}$$

$$p^4 + q^4 \equiv 1 - 4pq + 2p^2q^2 \tag{PQ5}$$

2 A quick guide to solving linear recurrence relations

2.1 Introduction

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

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

where $a_0, a_1, \ldots a_k$ 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 $x_0, x_1, \ldots x_{k-1}$ (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) $\theta_1, \theta_2, \ldots \theta_k$ are distinct, then the general solution of (R1) (where $g(n) \equiv 0$ for now) is :

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

and $c_1, c_2, \ldots c_k$ 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 $x_n = \kappa_n + \pi_n$.

2.2 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 :

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

2.3 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 $c_3 = 2$

2.4 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 $c_2 = -1$.

3 Grimmet and Welsh Chapter 1 Exercise 11

3.1 Question

Show that if a coin is tossed n times, then there are exactly

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

sequences of possible outcomes in which exactly k heads are obtained. If the coin is fair (so that heads and tails are equally likely on each toss), show that the probability of getting at least k heads is

$$\frac{1}{2^n}\sum_{r=k}^{n}\binom{n}{r}$$

3.2 Answer 1

A direct approach by enumerating the possibilities and trying to place k heads into n available positions :
The number of positions we can pick out of the n for the 1st head we place is n.
The number of positions we can pick out of the n for the 2nd head we place is $n-1$.
...
The number of positions we can pick out of the n for the kth head we place is $n-k+1$.
$\therefore$ The total number of permutations is $n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$
But, every permutation of the k heads has been included in the above, so we need to divide by $k!$:
Number of possible outcomes =

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

3.3 Answer 2

Use induction on n, where (clearly) $n \geq k$.
If $n = k$ then the number of possible outcomes is $\frac{n!}{k!(n-k)!} \equiv 1$ which is true.
If it is true for $n = t$ (where $t \geq k$) then number of possible outcomes is $\frac{t!}{k!(t-k)!}$.
For $t+1$ tosses, the number of possible outcomes is the total of those in which the first toss is a head, and those in which the first toss is a tail :

$$\begin{aligned}
&\frac{t!}{(k-1)!(t-k+1)!} + \frac{t!}{k!(t-k)!} \\
&= \frac{t!(k+t-k+1)}{k!(t-k+1)!} \\
&= \frac{(t+1)!}{((t+1)-k)!k!}
\end{aligned}$$

so the statement is therefore true for $n = t+1$, and therefore all $n \geq k$. Q.E.D.

4 Grimmet and Welsh Chapter 1 Problem 1

4.1 Question

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

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

4.2 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

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

4.3 Answer 1

A direct approach!
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.

4.4 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, using the partition theorem (F1.9) 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 :

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

4.5 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 $n \geq 0$.

4.6 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

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

5 Grimmet and Welsh Chapter 1 Exercise 27

5.1 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 $p = \frac{2}{3}$

5.2 Answer 1

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 $9\theta^2 - 3\theta - 2 = 0 \rightarrow (3\theta + 1)(3\theta - 2) = 0$, yielding the general solution :

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

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

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

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.

5.3 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!

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

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.

5.4 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,

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

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

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

6 Grimmet and Welsh Chapter 1 Problem 5

6.1 Question

Urn I has 4 white and 3 black balls. Urn II has 3 white and 7 black balls. An urn is selected at random and a ball is picked.

1. What is the probability that this is black?

2. If the ball is white, what is the probability that Urn I was selected?

6.2 Answer to part 1

$\mathbf{P}$(selected ball is black) =
$\mathbf{P}$(selected ball is black | Urn I was selected) $\times$ $\mathbf{P}$(Urn I was selected) $+$
$\mathbf{P}$(selected ball is black | Urn II was selected) $\times$ $\mathbf{P}$(Urn II was selected)

$$= \left(\frac{3}{7}\right)\left(\frac{1}{2}\right) + \left(\frac{7}{10}\right)\left(\frac{1}{2}\right) = \frac{79}{140}$$

6.3 Answer to part 2

$\mathbf{P}$(selected ball is white) $= 1 - \frac{79}{140} = \frac{61}{140}$
$\mathbf{P}$(selected ball is white | Urn I was selected) $= \frac{4}{7}$
From the definition of conditional probability (see (F1.6)) :

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

$$\therefore \mathbf{P}(A \mid B) = \mathbf{P}(B \mid A) \times \frac{\mathbf{P}(B)}{\mathbf{P}(A)} \text{ provided } \mathbf{P}(B) > 0 \text{ and } \mathbf{P}(A) > 0$$

So,

$$\mathbf{P}(\text{Urn I was selected} \mid \text{selected ball is white}) =$$

$$\mathbf{P}(\text{selected ball is white} \mid \text{Urn I was selected}) \times \frac{\mathbf{P}(\text{Urn I was selected})}{\mathbf{P}(\text{selected ball is white})}$$

$$= \frac{4}{7} \times \frac{\frac{1}{2}}{\frac{61}{140}} = \frac{40}{61}$$

7 Grimmet and Welsh Chapter 1 Problem 6

7.1 Question

A single card is removed from a deck of 52 cards. From the remainder we draw 2 cards at random and find that they are both spades.
What is the probability that the first card removed was also a spade?

7.2 General observation

Given the 'large' number of cards of each suit, we would still expect the answer to be quite close to $\frac{1}{4}$, but slightly lower since one of the possible spades was removed.

7.3 Answer

Firstly, we evaluate the probability of drawing 2 spades from the remaining 51 cards, conditioned on the first card.
Let $2\spadesuit$ be the event of drawing 2 spades, $\spadesuit_1$ be the event of the removed card being a spade, and $!\spadesuit_1$ be the event that the removed card is not a spade.

$$\mathbf{P}(2\spadesuit) = \mathbf{P}(2\spadesuit \mid \spadesuit_1) \times \mathbf{P}(\spadesuit_1) + \mathbf{P}(2\spadesuit \mid !\spadesuit_1) \times \mathbf{P}(!\spadesuit_1)$$

$$= \left(\frac{12}{51}\right)\left(\frac{11}{50}\right) \times \left(\frac{1}{4}\right) + \left(\frac{13}{51}\right)\left(\frac{12}{50}\right) \times \left(\frac{3}{4}\right) = \frac{33 + 117}{51 \times 50} = \frac{1}{17}$$

From the definition of conditional probability (see (F1.6)) :

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

$$\therefore \mathbf{P}(A \mid B) = \mathbf{P}(B \mid A) \times \frac{\mathbf{P}(B)}{\mathbf{P}(A)} \text{ provided } \mathbf{P}(B) > 0 \text{ and } \mathbf{P}(A) > 0$$

So,

$$\mathbf{P}(\spadesuit_1 \mid 2\spadesuit) = \mathbf{P}(2\spadesuit \mid \spadesuit_1) \times \left(\frac{\mathbf{P}(\spadesuit_1)}{\mathbf{P}(2\spadesuit)}\right)$$

$$= \frac{12}{51} \times \frac{11}{50} \times \frac{1}{4} \div \frac{1}{17} = \frac{12 \times 17 \times 11}{4 \times 51 \times 50} = \frac{11}{50}$$

8 Grimmet and Welsh Chapter 1 Problem 7

8.1 Question

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

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

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

8.3 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.

8.4 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.

8.5 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 :

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

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

8.6 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,

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

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

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

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

and we could try to determine the coefficient of x^n in $(1 + p^2x)^n(1 + q^2x)^n$
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!

9 Grimmet and Welsh Chapter 1 Problem 12b

9.1 Question

Part (a) of this question was to prove the formula (F1.5) :

$$\mathbf{P}(\bigcup_{1\le i\le n} A_i) = \sum_{1\le i\le n} \mathbf{P}(A_i) - \sum_{1\le i<j\le n} \mathbf{P}(A_i \cap A_j) + \ldots + (-1)^{n+1}\mathbf{P}(\bigcap_{1\le i\le n} A_i) \tag{10}$$

which we'll assume has been done!
One evening, a bemused lodge-porter tried to hang n keys on their n hooks, but only managed to hang them independently and at random. There was no limit to the number of keys which could be hung on any hook. Otherwise, or by using (a), find an expression for the probability that at least one key was hung on its own hook.
The following morning the porter was rebuked by the Bursar, so that in the evening he was careful to hang only one key on each hook. But he still only managed to hang them independently and at random. Find an expression for the probability that no key was then hung on its own hook.
Find the limits of both expressions as n tends to infinity.
You may assume that

$$e^x = \sum_{r=0}^{\infty} \frac{x^r}{r!} = \lim_{N\to\infty}\left(1+\frac{x}{N}\right)^N \tag{11}$$

for all real x.

9.2 Answer to first bit

The probability $\mathbf{P}_n$ that NO key was hung on its own hook is

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

as each key has $n-1$ positions out of n to be placed on to avoid being placed on its own hook.
So, the probability that at least one key was hung on its own hook is

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

Putting $x = -1$ in (11) we see that $\lim_{n\to\infty} 1 - \mathbf{P}_n = 1 - \frac{1}{e}$ which is approximately 0.632.

9.3 Answer to second bit

Let A_i be the event that key "i" is on hook "i". $(1 \le i \le n)$
The probability $\mathbf{P}_n$ that at least one key was hung on its own hook (with no duplicates allowed) is :

$$\mathbf{P}_n = \mathbf{P}\left(\bigcup_1^n A_i\right)$$

i.e. one or more of the events A_1, A_2, has occurred.
From part (a), this is equal to

$$\mathbf{P}(\bigcup_{1\le i\le n} A_i) = \sum_{1\le i\le n} \mathbf{P}(A_i) - \sum_{1\le i<j\le n} \mathbf{P}(A_i \cap A_j) + \ldots + (-1)^{n+1}\mathbf{P}(\bigcap_{1\le i\le n} A_i) \tag{12}$$

Looking specifically at $\sum_{1\le i<j\le n} \mathbf{P}(A_i \cap A_j)$, there are $\binom{n}{2}$ ways of choosing the values of i and j, and for each pair of values, key i has n possible hooks, and given that no multiples of keys are allowed on any hook, there are $n-1$ remaining possible hooks for key j.
So this term has value $\binom{n}{2}\frac{1}{n(n-1)}$ with similar arguments for the other terms.

$$\implies \mathbf{P}(\bigcup_{1\le i\le n} A_i) =$$

$$\binom{n}{1}\frac{1}{n} - \binom{n}{2}\frac{1}{n(n-1)} + \binom{n}{3}\frac{1}{n(n-1)(n-2)} - \cdots + (-1)^{n+1}\binom{n}{n}\frac{1}{n(n-1)(n-2)\ldots 1}$$

$$= \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n+1}\frac{1}{n!} \tag{13}$$

Again, putting $x = -1$ in (11) and substracting from 1, we see that (noting the slightly different definition of $\mathbf{P}_n$) : $\lim_{n\to\infty} \mathbf{P}_n = 1 - \frac{1}{e}$
And, as the question asked for the probability that no key was hung on its own hook, the answer(s) are just obtained by subtracting from 1.

9.4 Comment

The probabilities for at least one key on its own hook being $\sim 1 - e^{-1}$ for both scenarios might seem counter-intuitive. If, however, we use part (a) to solve the first question with multiple keys allowed on any hook, the only difference is that the denominators in (13) are powers of n rather than $n(n-1)(n-2)\ldots$
For the third term for instance, the expression will be

$$\binom{n}{3}\frac{1}{n^3} = \frac{1}{3!}\left((1-\frac{2}{n})(1-\frac{1}{n})(1-\frac{0}{n})\right) \tag{14}$$

With careful handling, the whole expression can still be shown to be $1 - \frac{1}{e} + \mathcal{O}\frac{1}{n}$ which has the same limit.

10 Grimmet and Welsh Chapter 1 Problem 13

10.1 Question

Two identical decks of cards, each containing N cards, are shuffled randomly. We say that a k-matching occurs if the two decks agree in exactly k places. Show that the probability that there is a k-matching is

$$\pi_k = \frac{1}{k!}\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{N-k}}{(N-k)!}\right) \tag{15}$$

for $k = 0, 1, 2, \ldots, N-1, N$. We note that $\pi_k \simeq (k!e)^{-1}$ for large N and fixed k. Such matching probabilities are used in testing departures from randomness in circumstances such as psychological tests and wine-tasting competitions. (The convention is that $0! = 1$.)

10.2 Answer

π_k is the probability of exactly k matches, which arises if there are k matches amongst the N cards and the remaining $N-k$ are not matched.
Now, k matches can occur in $\binom{N}{k}$ ways, each with probability $\frac{1}{N}\frac{1}{N-1}\cdots\frac{1}{N-k+1}$.
From the answers to problem 12b, we also know that the probability of no matches amongst the $N-k$ cards *given that the other* k *cards match* (so the $N-k$ cards from each deck contain the same values amongst them), is :

$$= 1 - \left(\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{N-k+1}\frac{1}{(N-k)!}\right)$$

Using the fact (see (F1.6)) that $\mathbf{P}(A \cap B) = \mathbf{P}(B)\mathbf{P}(A \mid B)$, we obtain

$$\pi_k = \frac{1}{N}\frac{1}{N-1}\cdots\frac{1}{N-k+1}\binom{N}{k}\left[1 - \left(\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{N-k+1}\frac{1}{(N-k)!}\right)\right]$$

$$= \frac{1}{k!}\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{N-k}\frac{1}{(N-k)!}\right]$$

Q.E.D.

10.3 Comment

As noted in the question, $\pi_k \simeq (k!e)^{-1}$ for large N and fixed k.
This gives the interesting result that $\pi_0 \simeq \pi_1 \simeq \frac{1}{e}$ for even quite small N.

There are a couple of other checks we can do, firstly setting $k = N$. In this case the expression instantly collapses to $\frac{1}{N!}$ which is as we expect.

Also, with no justification whatsoever that it is valid, we note that the sum of the approximations of π_k for $0 \leq k \leq N$ is

$$\sum_{k=0}^{N} \pi_k \simeq \frac{1}{e}\left[\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^N\frac{1}{(N)!}\right] \simeq 1$$

11 Grimmet and Welsh Chapter 1 Problem 14

11.1 Question

The buses which stop at the end of my road do not keep to the timetable. They should run every quarter hour, at 8:30, 8:45, 9:00 ... but in fact each bus is either five minutes early or five minutes late, the two possibilities being equally likely and the busses independent.
Other people arrive at the bus stop in such a way that t minutes after the departure of one bus, the probability that no one is waiting for the next one is $e^{\frac{-t}{5}}$.
(a) What is the probability that no one is waiting at 9:00?
(b) One day I come to the bus stop at 9:00 and find no one there; show that the chances are more than four to one that I have missed the nine o'clock bus. (You may use hthe good approximation $e^3 \approx 20$)

11.2 Answer to part (a)

Let $B_{h:mm}$ be the event that a bus arrives at time $h : mm$.
There are three possibilities for the times of the most recent bus prior to 9:00, namely :
(i) 8:55 - The 9:00 bus was 5 minutes early,
(ii) 8:50 - The 8:45 bus was 5 minutes late and the 9:00 bus will be 5 minutes late,
(iii) 8:40 - The 8:45 bus was 5 minutes early and the 9:00 bus will be 5 minutes late.

$$\begin{aligned}
&\therefore \mathbf{P}(\text{no one is waiting at 9:00}) \\
&\qquad = \mathbf{P}(B_{8:55})e^{-1} + \mathbf{P}(B_{8:50})\mathbf{P}(B_{9:05})e^{-2} + \mathbf{P}(B_{8:40})\mathbf{P}(B_{9:05})e^{-4} \\
&\qquad = \frac{1}{4}(2e^{-1} + e^{-2} + e^{-4})
\end{aligned}$$

11.3 Answer to part (b)

Note that "missed the nine o'clock bus" means that the bus must have arrived at 8:55 (with probability $\frac{1}{2}$ ignoring any other constraints).

$$\begin{aligned}
&\mathbf{P}(B_{8:55} \mid \text{no one is waiting at 9:00}) \\
&\quad = \mathbf{P}(\text{no one is waiting at 9:00} \mid B_{8:55}) \times \frac{\mathbf{P}(B_{8:55})}{\mathbf{P}(\text{no one is waiting at 9:00})} \\
&\quad = e^{-1} \times \frac{\frac{1}{2}}{\frac{1}{4}(2e^{-1} + e^{-2} + e^{-4})} = \frac{2e^3}{2e^3 + e^2 + 1} \\
&\quad \approx \frac{40}{49} > \frac{4}{5}
\end{aligned}$$

Q.E.D.

12 Grimmet and Welsh Chapter 1 Problem 15

12.1 Question

A coin is tossed repeatedly; on each toss a head is shown with probability p, a tail with probability $1-p$. All tosses are mutually independent. Let E denote the event that the first run of r successive heads occurs earlier than the first run of s successive tails. Let A denote the outcome of the first toss. Show that

$$\mathbf{P}(E \mid A = \text{head}) = p^{r-1} + (1-p^{r-1})\mathbf{P}(E \mid A = \text{tail})$$

Find a similar expression for $\mathbf{P}(E \mid A = \text{head})$ and hence find $\mathbf{P}(E)$.

12.2 Answer

If A is a head, there are two options which can contribute to E occurring :

(i) The next $r-1$ tosses are all heads, with probability p^{r-1} - which means E has occured.

(ii) The next $r-1$ tosses are not all heads (with probability $1-p^{r-1}$), so a tail appears, at which point the 'counts' of r and s are reset to the state after the first toss was a tail.

$$\therefore \mathbf{P}(E \mid A = \text{head}) = p^{r-1} + (1-p^{r-1})\mathbf{P}(E \mid A = \text{tail})$$

If A is a tail, there is one option (since a run of s tails means that E does not occur) :

(i) The next $s-1$ tosses are not all tails (with probability $1-(1-p)^{s-1}$), so a head appears, at which point the 'counts' of r and s are reset to the state after the first toss was a head.

$$\therefore \mathbf{P}(E \mid A = \text{tail}) = \left(1-(1-p)^{s-1}\right)\mathbf{P}(E \mid A = \text{head}) \tag{16}$$

Eliminating $\mathbf{P}(E \mid A = \text{tail})$,

$$\mathbf{P}(E \mid A = \text{head}) = \frac{p^{r-1}}{1-(1-p^{r-1})\left(1-(1-p)^{s-1}\right)}$$

and now using (16) :

$$\mathbf{P}(E \mid A = \text{tail}) = \frac{p^{r-1}\left(1-(1-p)^{s-1}\right)}{1-(1-p^{r-1})\left(1-(1-p)^{s-1}\right)}$$

But $\mathbf{P}(E) = \mathbf{P}(E \mid A = \text{head})p + \mathbf{P}(E \mid A = \text{tail})(1-p)$, so

$$\begin{aligned}\mathbf{P}(E) &= \frac{p^r + (1-p)p^{r-1}\left(1-(1-p)^{s-1}\right)}{1-(1-p^{r-1})\left(1-(1-p)^{s-1}\right)} \\ &= \frac{p^{r-1}(1-(1-p)^s)}{p^{r-1}+(1-p)^{s-1}-p^{r-1}(1-p)^{s-1}}\end{aligned}$$

12.3 Comment

It is a messy expression, but we can do one reasonability check, namely that if $r = s$ and $p = (1-p) = \frac{1}{2}$:

$$\mathbf{P}(E) = \frac{\frac{1}{2}^{r-1}(1-\frac{1}{2}^r)}{\frac{1}{2}^{r-1}+\frac{1}{2}^{r-1}-\frac{1}{2}^{r-1}\frac{1}{2}^{r-1}} = \frac{1}{2}$$

as expected!

13 Runs of either heads or tails

13.1 Question

A biased coin shows heads with probability $p = 1 - q$ whenever it is tossed.
What is the probability that in n tosses, there is no run of length R of *either* heads or tails? $(n \geq 1)$

13.2 Comment 1

I suspect that a general solution, even establishing a general recurrence relation for any R is quite tricky. All I have done is to evaluate some particular cases, and presume that the same method can applied for any larger specific value of R.

13.3 Answer for $R = 2$

The question for $R = 2$ is actually very easy, since after the first toss, all subsequent values must alternate between heads and tails to avoid a run of 2.
For even values of n , $v_n = 2p^{n/2}q^{n/2}$, and for odd values of n,
$v_n = p^{(n+1)/2}q^{(n-1)/2} + p^{(n-1)/2}q^{(n+1)/2} = p^{(n-1)/2}q^{(n-1)/2}$, however this is a useful example to cut our teeth on!
Let A_j represent the event that, in j tosses, no pair of heads or tails occur successively and let $v_j := \mathbf{P}(A_j)$.
Also, define H_j as $\mathbf{P}(A_j \mid \text{first toss is a head})$ and T_j as $\mathbf{P}(A_j \mid \text{first toss is a tail})$.
Conditioning on the result of the first toss for v_{n+2},

$$v_{n+2} = \mathbf{P}(A_{n+2}) = H_{n+2}p + T_{n+2}q$$

But, H_{n+2} is the probability of no pair occurring in $n + 2$ tosses given a head on the first toss, which means that the subsequent toss can only be a tail (with probability q) and the 'state' is then reset to that in which there are $n + 1$ tosses remaining with the "first" toss being a tail. So

$$H_{n+2} = qT_{n+1} \qquad \text{and similarly } T_{n+2} = pH_{n+1} \tag{17}$$

which gives :

$$v_{n+2} = pq(T_{n+1} + H_{n+1})$$

Using (17) again, but with n reduced by 1, this leads to

$$\begin{aligned} v_{n+2} &= pq(pH_n + qT_{n+1}) \\ &= pqv_n \end{aligned}$$

Also, we can determine by looking at the possibilities, that $v_1 = 1$ and $v_2 = pq + qp = 2pq$.
From this, we can directly obtain the solution mentioned at the beginning of this section.
Alternatively, we can solve the recurrence relation :
The auxiliary equation is $\theta^2 - pq = 0 \implies \theta = \pm\sqrt{pq}$, yielding the general solution :

$$v_n = C(-\sqrt{pq})^n + D(\sqrt{pq})^n \text{ for some } C \text{ and } D.$$

From the initial conditions,

$$\begin{aligned} v_1 = 1 &\implies 1 = (-C + D)\sqrt{pq} \\ v_2 = 2pq &\implies 2pq = (C + D)pq \implies C + D = 2 \\ &\implies C = 1 - \frac{1}{2\sqrt{pq}} \text{ and } D = 1 + \frac{1}{2\sqrt{pq}} \end{aligned}$$

So :

$$v_n = \left(1 - \frac{1}{2\sqrt{pq}}\right)(-\sqrt{pq})^n + \left(1 + \frac{1}{2\sqrt{pq}}\right)(\sqrt{pq})^n$$

This looks quite unusual, but for even and odd values of n, different pairs of terms cancel each other out to give the result we had before.
Note that for $p = q = \frac{1}{2}$, the values of v_1, v_2, v_3, v_4 etc. are 1, $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}$ ....

13.4 Answer for $R = 3$

This time the solution is not obvious (to me anyway!), but we can try the same approach.
Let A_j represent the event that, in j tosses, no run of three heads or tails occur successively and let $v_j := \mathbf{P}(A_j)$.
Also, define H_j as $\mathbf{P}(A_j \mid \text{first toss is a head})$ and T_j as $\mathbf{P}(A_j \mid \text{first toss is a tail})$.
Conditioning on the result of the first toss for v_{n+4},

$$v_{n+4} = \mathbf{P}(A_{n+4}) = H_{n+4}p + T_{n+4}q$$

But, H_{n+4} is the probability of no run of three occurring in $n + 4$ tosses given a head on the first toss, which means that either :

(i) The next toss is a tail (with probability q) and the 'state' is then reset to that in which there are $n + 3$ tosses remaining with the “first” toss being a tail

(ii) The next toss is a head, followed by a tail (with probability pq) and the 'state' is then reset to that in which there are $n + 2$ tosses remaining with the “first” toss being a tail

So

$$H_{n+4} = qT_{n+3} + pqT_{n+2} \tag{18}$$

Similarly :

$$T_{n+4} = pH_{n+3} + qpH_{n+2} \tag{19}$$

which gives after substituting :

$$v_{n+4} = pqT_{n+3} + p^2qT_{n+2} + pqH_{n+3} + q^2pH_{n+2}$$

As for the $R = 2$ solution, using (18) and (19) again, but with n reduced by 1, and substituting a further time, this leads to

$$\begin{aligned}
v_{n+4} &= pq\,(pH_{n+2} + qpH_{n+1}) + p^2q\,(pH_{n+1} + qpH_n) \\
&\qquad + pq\,(qT_{n+2} + qpT_{n+1}) + q^2p\,(qT_{n+1} + qpT_n) \\
&= pq\left(pH_{n+2} + qT_{n+2} + (qp + p^2)H_{n+1} + (qp + q^2)T_{n+1} + qp^2H_n + pq^2T_n\right)
\end{aligned}$$

and noting that $qp + p^2 \equiv p$, and $qp + q^2 \equiv q$:

$$v_{n+4} = pq\,(v_{n+2} + v_{n+1} + qpv_n)$$

To get the initial conditions, with some slightly laborious counting, we can determine that :

$$\begin{aligned}
v_1 &= 1 \\
v_2 &= 1 \\
v_3 &= 1 - p^3 - q^3 \equiv 3pq \\
v_4 &= 2p^3q + 6p^2q^2 + 2pq^3 \equiv 2pq(p^2 + 3pq + q^2) \equiv 2pq(1 + pq)
\end{aligned}$$

Attempting to solve this we get the auxiliary equation is $\theta^4 - pq(\theta^2 + \theta + pq) = 0$ which does not factorise for θ in terms of p. We can still derive the values for any given values of p however.

Below is a table of probabilities for the first few values of n for $p = 0.4$ and $p = 0.5$ from a spreadsheet using the recurrence relationship above. As it is easy to make mistakes deriving formulae, it seems a good idea to verify results by using a separate mechanism - in this case, a simple Java program which simply 'makes', for each value of n, n tosses of a coin millions of times and counts the proportion which don't contain a run of three consecutive heads or tails.
As you can see, the results suggest that the recurrence relation and initial conditions matches the results from the 'real' sampling and are presumably correct!

	p = 0.5				**p = 0.4**		
n	v_n	**Ratio**	**Java**		v_n	**Ratio**	**Java**
1	1		1.0000000		1		1.0000000
2	1	1	1.0000000		1	1	1.0000000
3	0.75	0.75	0.7499893		0.72	0.72	0.7199326
4	0.625	0.83333333	0.6250490		0.5952	0.82666667	0.5951957
5	0.5	0.8	0.5000507		0.4704	0.79032258	0.4704234
6	0.40625	0.8125	0.4063904		0.373248	0.79346938	0.3732993
7	0.328125	0.80769230	0.3282143		0.297216	0.79629629	0.2971779
8	0.265625	0.80952381	0.2657503		0.23675904	0.79658914	0.2367888
9	0.21484375	0.80882352	0.2149047		0.1880064	0.79408330	0.1878943
10	0.173828125	0.80909090	0.1738321		0.149653094	0.796	0.1496198
11	0.140625	0.80898876	0.1406627		0.119063347	0.79559562	0.1191090
12	0.113769531	0.80902777	0.1138066		0.094675599	0.7951699	0.0946223
13	0.092041016	0.80901287	0.0920047		0.075321115	0.79557050	0.0753025
14	0.074462891	0.80901856	0.0745270		0.059917365	0.79549228	0.0599458
15	0.060241699	0.80901639	0.0602841		0.04765726	0.79538310	0.0476805

(The first 4 values of v_n are 'hardcoded' from the initial conditions, the subsequent values are derived using the recurrence relation)
For $p = 0.5$, the auxiliary equation turns out to be $\theta^4 - \frac{1}{4}\theta^2 - \frac{1}{4}\theta - \frac{1}{16} = 0$ which has the four solutions, 0.8090169943749475, -0.30901699437494745, $-0.25 + 0.43301270189221935i$, $-0.25 - 0.43301270189221935i$.
The complex solutions would always have constants which are designed to cancel out any lingering imaginary components, but regardless, the "largest" solution , i.e. the one whose powers decreases slowest to 0 is 0.8090169943749475 which happens in this case to be $\frac{\sqrt{5}+1}{4}$.
We can see from the above table that the successive ratios appear to be converging to this value. Similar comments apply to other values of p.

13.5 Answer for $R = 4$

This is only slightly more complex than the $R = 3$ case, mainly in the areas of larger formulae and more effort getting the initial conditions. Feel free to skip this section if uninterested, as there's nothing essentially new here apart from perhaps confirming some of the patterns we might have seen from the previous two solutions!
Let A_j represent the event that, in j tosses, no run of four heads or tails occur successively and let $v_j := \mathbf{P}(A_j)$.
Also, define H_j as $\mathbf{P}(A_j \mid \text{first toss is a head})$ and T_j as $\mathbf{P}(A_j \mid \text{first toss is a tail})$.
Conditioning on the result of the first toss for v_{n+6},

$$v_{n+6} = \mathbf{P}(A_{n+6}) = H_{n+6}p + T_{n+6}q$$

But, H_{n+6} is the probability of no run of four occurring in $n + 6$ tosses given a head on the first toss, which means that either :

(i) The next toss is a tail (with probability q) and the 'state' is then reset to that in which there are $n + 5$ tosses remaining with the "first" toss being a tail

(ii) The next toss is a head, followed by a tail (with probability pq) and the 'state' is then reset to that in which there are $n+4$ tosses remaining with the "first" toss being a tail

(iii) The next two tosses are heads, followed by a tail (with probability p^2q) and the 'state' is then reset to that in which there are $n+3$ tosses remaining with the "first" toss being a tail

So
$$H_{n+6} = qT_{n+5} + pqT_{n+4} + p^2qt_{n+3} \tag{20}$$
Similarly :
$$T_{n+6} = pH_{n+5} + qpH_{n+4} + q^2pH_{n+3} \tag{21}$$
which gives after substituting :

$$v_{n+6} = pqT_{n+5} + p^2qT_{n+4} + p^3qT_{n+3} + pqH_{n+5} + q^2pH_{n+4} + q^3pH_{n+3}$$

And again, using (20) and (21), but with n reduced by 1, and substituting a further time, (and noting that $qp + p^2 \equiv p$ and $pq + q^2 \equiv q$) this leads to :

$$\begin{aligned}
v_{n+4} &= pq\left(pH_{n+4} + qpH_{n+3} + q^2pH_{n+2}\right) + p^2q\left(pH_{n+3} + qpH_{n+2} + q^2pH_{n+1}\right) \\
&\quad + p^3q\left(pH_{n+2} + qpH_{n+1} + q^2pH_n\right) + pq\left(qT_{n+4} + pqT_{n+3} + p^2qT_{n+2}\right) \\
&\quad + pq^2\left(qT_{n+3} + pqT_{n+2} + p^2qT_{n+1}\right) + pq^3\left(qT_{n+2} + pqT_{n+1} + p^2qT_n\right) \\
&= pq\Big\{pH_{n+4} + qT_{n+4} + (qp+p^2)H_{n+3} + (pq+q^2)T_{n+3} \\
&\quad + (q^2p + qp^2 + p^3)H_{n+2} + (p^2q + pq^2 + q^3)T_{n+2} \\
&\quad + (q^2p^2 + qp^3)H_{n+1} + (p^2q^2 + pq^3)T_{n+1} + q^2p^3H_n + p^2q^3T_n\Big\} \\
&= pq\left(v_{n+4} + v_{n+3} + (1-pq)v_{n+2} + pqv_{n+1} + p^2q^2v_n\right)
\end{aligned}$$

Phew! To get the initial conditions, we need careful counting after which we can determine that (after some simplification) :

$$\begin{aligned}
&v_1 = 1 \\
&v_2 = 1 \\
&v_3 = 1 \\
&v_4 = 1 - p^4 - q^4 \equiv 4p^3q + 6p^2q^2 + 4pq^3 \equiv 4pq - 2p^2q^2 \\
&v_5 = 3p^4q + 10p^3q^2 + 10p^2q^3 + 3pq^4 \equiv pq(3p^3 + 3q^3 + 10(p^2q + q^2p)) \equiv 3pq + p^2q^2 \\
&v_6 = 2p^5q + 12p^4q^2 + 20p^3q^3 + 12p^2q^4 + 2pq^5 \equiv 2pq + 4p^2q^2
\end{aligned}$$

As before, below is a table of probabilities for the first few values of n for $p = 0.4$ and $p = 0.5$ from a spreadsheet using the recurrence relationship above including the results from a 'brute force' Java program to lend some validity to the formulae.
Again, the results suggest that the recurrence relation and initial conditions matches the results from the 'real' sampling and are presumably correct!

	p = 0.5				**p = 0.4**		
n	v_n	**Ratio**	**Java**		v_n	**Ratio**	**Java**
1	1		1.0000000		1		1.0000000
2	1	1	1.0000000		1	1	1.0000000
3	1	1	1.0000000		1	1	1.0000000
4	0.875	0.875	0.8749609		0.8448	0.8448	0.8447761
5	0.8125	0.92857142	0.8126014		0.7776	0.92045454	0.7776422
6	0.75	0.92307692	0.7500013		0.7104	0.91358024	0.7105294
7	0.6875	0.91666666	0.6876102		0.6432	0.90540540	0.6431889
8	0.6328125	0.92045454	0.6329641		0.58263552	0.90583880	0.5825617
9	0.58203125	0.91975308	0.5820986		0.52918272	0.90825688	0.5290458
10	0.53515625	0.91946308	0.5351430		0.48024576	0.90752351	0.4801654
11	0.4921875	0.91970802	0.4922160		0.43582464	0.90750335	0.4357943
12	0.452636719	0.91964285	0.4526049		0.395404444	0.90725582	0.3954265
13	0.416259766	0.91963322	0.4162215		0.358831227	0.90750428	0.3588773
14	0.3828125	0.91964809	0.3828548		0.325627085	0.90746585	0.3255918
15	0.352050781	0.91964285	0.3521412		0.295488553	0.90744464	0.2955003

(The first 6 values of v_n are 'hardcoded' from the initial conditions, the subsequent values are derived using the recurrence relation)
For $p = 0.5$, the auxiliary equation turns out to be $\theta^6 - \frac{1}{4}\theta^4 - \frac{1}{4}\theta^3 - \frac{3}{16}\theta^2 - \frac{1}{16}\theta - \frac{1}{64} = 0$ which has the six solutions, $\pm 0.5i$, -0.5, 0.9196433776070806, $-0.2098216888035 \pm 0.3031453646036i$.
Again, the "largest" solution , i.e. the one whose powers decreases slowest to 0 is 0.9196433776070806 and we can see from the above table that the successive ratios appear to be converging to this value. Similar comments apply to other values of p.

13.6 Comments on solutions

There are some features which are worth noting and probably apply to all values of R.
Firstly, for a given value of R, the final recurrence relationship is of the form $v_{n+2R} = C_{n+2R-1}v_{n+2R-1} + C_{n+2R-2}v_{n+2R-2} + \cdots + C_n v_n$. I suspect that C_{n+2R-1} is always 0.
Secondly, each approach seems to require the double substitution of the formulae for H_i and T_i after which 'magically' allows us to rewrite in terms of v_i.
Thirdly, it is interesting that every recurrence relation and initial condition can be formulated in terms of pq and no other functions of p or q. This is probably because of (a) the obvious symmetry between p and q in the problem, and (b) that many symmetrical functions of 2 variables can often be rewritten as functions of the product and sum of those variables. Since $p + q \equiv 1$, we are only left with the product pq.