

# 5. PROBABILITY AND THE BINOMIAL THEOREM

## §5.1. Numbers of Arrangements

In how many ways can we arrange  $n$  things? If  $n = 3$  and the things are the letters A, B, C then the arrangements are: ABC, ACB, BAC, BCA, CAB, CBA.

There are 3 letters that can occupy the first place, 2 that can go in second place and 1 that can occupy the remaining place. That is,  $3 \times 2 \times 1 = 6$ . Note that we didn't have the list the arrangements and then count them. We have a system for working out the number without a list.

**Example 1:** In how many ways can the 26 letters of the alphabet be arranged?

**Solution:** The task of listing all the arrangements would be formidable! Yet we can determine their number, on the above principle, as  $26 \times 25 \times 24 \times \dots \times 3 \times 2 \times 1$ . Using a calculator you can determine that this number is just over  $4 \times 10^{26}$ , which is 4 followed by 26 zeros. The fact that the power of 10 is 26, the number of letters, is a bit of a coincidence.

For every positive integer  $n$  we define  $n!$  (pronounced  $n$  **factorial**) to be  $n(n - 1) \dots$  3.2.1.

One group of people who are very familiar with factorials are bell-ringers. Those are people who go up into a bell tower in a cathedral and pull the ropes. By the way don't ever ask a bell-ringer "are you a campanologist?". You might be showing off your erudite learning, but you will reveal that you aren't a bell-ringer yourself. Bell-ringers always call themselves 'bell-ringers'.

If ever your hear a melody floating down from a bell-tower you can be sure that it's someone at a keyboard operating the bells. Bell-ringing in the English tradition consists of ringing patterns, not tunes. That's not by choice but because of the physics. While carillons, operate with the bells being struck by hammers, a bell in the English method of ringing has to rotate a complete 360 degrees when its rope is pulled. The sound comes two thirds of the way round when the clapper, inside the bell, catches up with the bell itself.

As a consequence it takes about 2 seconds before the same bell can be sounded again. Imagine trying to play "Twink.....le.....twink.....le.....little.....star" on a set of bells!

The fundamental principle of English bell-ringing is that each bell rings in a certain order. Then they all ring in a different order, and so on. And the convention is that no arrangement is to be repeated. There's no practical, or even aesthetic, reason for this. If you listened to a peal of

bells, you're hardly likely to notice that the same arrangement is repeated half an hour later. But that's how it's always been done.

So, ringing on 8 bells there would be  $8!$  different arrangements. To go through all of these arrangements would take 24 hours! No-one ever does this. But there's a convention that the tenor bell, the largest, always rings last each time. This acts as a sort of punctuation mark after each of the other 7 have rung. So a full peal of bells consists of ringing the other 7 bells in all  $7!$  ways. As every bell-ringer knows,  $7! = 5040$  and a full peal takes about 3 hours, which is quite a feat!. I have only ever rung a quarter peal, which takes about 45 minutes.

How does one guarantee no repetitions? The answer is that there are certain methods. each involving a few simple rules, that can be proved on pencil and paper to contain no repetitions.

## **§5.2. Numbers of Choices**

How many ways are there of choosing 2 things out of 5? We might have 5 letters A, B, C, D, E and BD will be one such choice. The number depends on whether we consider BD and DB as the same choice or different choices.

If we want the choices in a specific order then they are:  
AB, AC, AD, AE, BA, BC, BD, BE, CA, CB,  
CD, CE, DA, DB, DC, DE, EA, EB, EC, ED.

There are altogether 20 choices. For each of the 5 possibilities for the first letter, there are 4 choices for the second.

But if the order of the letters does not matter we have only half of these:

AB, AC, AD, AE, BC, BD, BE, CD, CE, DE.

We divide by 2, because each choice will appear twice in the first list.

**Theorem 1:** The number of ways of choosing  $r$  things from  $n$  is:

(1)  ${}^n P_r = n(n-1)(n-2) \dots (n-r+1) = \frac{n!}{(n-r)!}$  if order matters (**permutations**) and

(2)  ${}^n C_r = \frac{n(n-1)(n-2) \dots (n-r+1)}{r(r-1)(r-2) \dots 3 \cdot 2 \cdot 1} = \frac{n!}{r!(n-r)!}$  if

order doesn't matter (**combinations**).

**Proof:**

(1) If the order matters there are  $n$  possibilities for the first choice.

For each of these there are  $n-1$  possibilities for the second choice, making  $n(n-1)$  possibilities for the first two choices.

For each of these there are  $n-2$  possibilities for the third choice, making  $n(n-1)(n-2)$  possibilities for the first three choices. And so on.

(2) If the choice doesn't matter each choice of  $r$  things can be arranged in  $r!$  ways and so we must divide by this number.

We denote the number of choices of  $r$  things from  $n$ , where the order matters, as  ${}^n\mathbf{P}_r$ . Where the order doesn't matter we use the symbol  ${}^n\mathbf{C}_r$ , or more usually, by the symbol  $\binom{n}{r}$ .

One way of looking at choosing  $r$  things from  $n$  is to think of it as separating the  $n$  things into two subsets, those that are accepted and those that are rejected.

**Example 2:** You have 24 students and you have 2 sports activities – tennis and running. As there are only 2 courts available you



have to choose 8 students to play tennis and the rest will have to go running. In how many ways can you make such a choice?

**Solution:** We have to separate the 24 students into two groups – the tennis group and the running group. It doesn't matter which group we choose first, as long as there are 8 in the tennis group and 16 in the running group. We could choose 8 students from the 24 to play tennis. The rest will run.

There are  $\binom{24}{8} = \frac{24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 19 \cdot 18 \cdot 17}{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$  possible choices.

If we cancel as far as possible we get:

$$\frac{3 \cdot 23 \cdot 11 \cdot 19 \cdot 3 \cdot 17}{1} = 735471.$$

Now we could have chosen the runners first, and the rest will play tennis.

There are:

$$\binom{24}{16} = \frac{24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

possible choices.

Now notice that we have the factor

$$16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9$$

in both the numerator and denominator. If we cancel these

we get  $\frac{24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 19 \cdot 18 \cdot 17}{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$  as before.

**Theorem 2:**  $\binom{n}{r} = \binom{n}{n-r}$  for all  $r, n$ .

**Proof:** The LHS is the number of ways of choosing  $r$  things and rejecting the other  $n - r$  while the RHS is the number of ways of choosing the  $n - r$  things and rejecting the rest. Both involve dividing the  $n$  things into a subset of size  $r$  and a subset of size  $n - r$ . 🙌😊

More generally, if we have  $n$  things and we wish to separate them into  $k$  subsets of sizes  $r_1, r_2, \dots, r_k$ , where  $r_1 + r_2 + \dots + r_k = n$ , the number of ways of doing this is:

$\frac{n!}{r_1! r_2! \dots r_k!}$ . This is called a **multinomial coefficient**.

The special case, where  $k = 2$ , is  $\frac{n!}{r! (n-r)!} = \binom{n}{r}$  called a **binomial coefficient** (for reasons that will become clear later).

**Theorem 3:** For all  $n, r$ :

$$(1) \binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1};$$

$$(2) \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = 2^n.$$

**Proof:**

(1) Suppose we have to choose  $r$  things from  $n+1$ , say  $r$  numbers from  $\{0, 1, 2, \dots, n\}$ .

Such a choice might include 0 or it might exclude 0.

How many exclude 0? In this case we have to choose all  $r$  of them from  $\{1, 2, \dots, n\}$  and the number of such choices is  $\binom{n}{r}$ .

How many include 0? In this case we have to choose the remaining  $r-1$  things from

$\{1, 2, \dots, n\}$  and the number of such choices is  $\binom{n}{r-1}$ .

Every choice must either include or exclude 0 (and no choice can do both) so the total number of choices is:

$$\binom{n}{r} + \binom{n}{r-1}.$$

(2) Suppose we have to choose a subset from  $\{1, 2, \dots, n\}$ .

The size of the subset can range from 0 (by choosing nothing) to  $n$  (by choosing everything). The number of choices, where the size is  $r$ , is  $\binom{n}{r}$ . Hence the total number

of choices is  $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n}$ .

But in choosing a subset from  $\{1, 2, \dots, n\}$  there are 2 possibilities for each of these numbers – either it is in the subset or it is not. So the total number of subsets is  $2^n$ .



**Example 3:** In how many ways can 8 people be seated around a circular dining table?

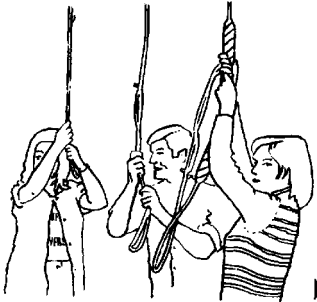
**Solution:** Here we are assuming that it makes no difference where each person actually sits, just who sits on the left and right of each person. So one person can be seated anywhere and the remaining 7 people can then be seated relative to that person in  $7!$  ( $= 5040$ ) ways.

**Example 4:** In how many ways can eight different coloured beads be put onto a necklace.

**Solution:** The assumption here is that the beads can be slid around the circle and that the necklace can be turned

over. So the answer this time will be a half of  $7!$  ( $= 2520$ ) ways.

**Example 5:** A peal of bells is a way of ringing all the bells in some order, then ringing them in a different order, and so on. The rules of bell-ringing insist that no permutation may be repeated and a full peal consists of every permutation being rung. A quarter peal is a methodical way in which a quarter of all the permutations



are rung. However a convention insists that the heaviest bell, called the tenor bell, always rings last in every permutation, so with 8 bells a full peal will consist of each bell being rung  $7!$  times. Typically the interval between a bell ringing one time and the next is about 2 seconds. How long would it take to ring a quarter peal?



**Solution:** A quarter peal consist of each bell being rung  $\frac{7!}{4}$  ( $= 1260$ ) times with an interval of about 2 seconds.

Hence the number of minutes to ring a quarter peal is about  $\frac{1260 \times 2}{60} = 42$  minutes.

### §5.3. Pascal's Triangle

There's a rather interesting way to list the binomial coefficients, called Pascal's Triangle. Start with 1 on the top. Underneath write "1 1". So far we have  $\begin{matrix} 1 \\ 1 & 1 \end{matrix}$ . Now write 1's diagonally to the left and the right and under each pair, between them, write the sum of those two numbers.

				1									
				1		1							
			1		2		1						
		1		3		3		1					
	1		4		6		4		1				
	1	5		10		10		5	1				
1		6		15		20		15		6		1	
1	7		21		35		35		21		7		1

The  $n$ 'th row of this triangle give the binomial coefficients  $\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n}$ .

### §5.4. The Binomial Theorem

It would be possible to expand  $(a + b)^6$  by putting  $(a + b)^2 = a^2 + 2ab + b^2$  and then expanding  $(a + b)^4 = (a^2 + 2ab + b^2)(a^2 + 2ab + b^2)$  and finally multiplying the answer by  $a^2 + 2ab + b^2$  again. But this would be very tedious.

Imagine expanding

$$(a + b)(a + b)(a + b)(a + b)(a + b)(a + b).$$

Each term will consist of 6 factors, one chosen from each of the above factors. Sometimes we may choose an  $a$ , and sometimes a  $b$ . Suppose we choose  $a$  from the first factor,  $b$  from the second and third,  $a$  from the fourth factor and  $b$  from the last two factors. This will give the term  $abbabb$ , which simplifies to  $a^2b^4$ .

There'll be other terms that simplify to  $a^2b^4$ . How many? As many ways as there are of choosing 2 factors, from the 6, where the choice is an  $a$ . There are  $\binom{6}{2} = 15$  such choices. So altogether the coefficient of  $a^2b^4$  will be 10.

Now the individual terms of  $(a + b)^6$  will all have the form  $a^r b^{6-r}$  corresponding to a choice of  $r$  letter  $a$ 's and

$6 - r$  letter  $b$ 's. There will be  $\binom{6}{r}$  such terms which can be combined as  $\binom{6}{r} a^r b^{6-r}$ .

Finally,  $r$  can range from 0 to 6. When  $r = 0$  we choose no  $a$ 's and all  $b$ 's.

When  $r = n$  we choose all  $a$ 's and no  $b$ 's.

There will be just one choice for each, so the coefficients of  $a^6$  and  $b^6$  will each be 1.

$$\text{So } (a + b)^6 = a^6 + \binom{6}{1} a^5 b + \binom{6}{2} a^4 b^2 + \binom{6}{3} a^3 b^3 \\ + \binom{6}{4} a^2 b^4 + \binom{6}{5} a b^5 + b^6.$$

Evaluating the binomial coefficients this becomes:

$$a^6 + 6a^5 b + 15a^4 b^2 + 20a^3 b^3 + 15a^2 b^4 + 6ab^5 + b^6.$$

Notice that these coefficients can be read off from the 6<sup>th</sup> row of Pascal's Triangle.

**Theorem 4 (BINOMIAL THEOREM):** For all real numbers  $a$ ,  $b$  and all positive integers  $n$ :

$$(a + b)^n = a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \dots + \binom{n}{r} a^{n-r} b^r \\ + \dots + \binom{n}{n-1} a b^{n-1} + b^n.$$

**Proof:** The terms in  $a^{n-r} b^r$  will result from  $n - r$  choices of  $a$  and  $r$  choices of  $b$  from  $(a + b)^n$ .

There will be  $\binom{n}{r}$  such choices and hence the coefficient of  $a^{n-r} b^r$  is  $\binom{n}{r}$ . 🙌😊

Notice the pattern of the RHS. The powers of  $a$  start at  $a^n$  and lose a factor of  $a$  at each step, while the powers of  $b$  pick up an extra factor of  $b$  at each step. The coefficient

is  $\binom{n}{r}$  where  $r$  is the number of  $b$ 's. But since  $\binom{n}{r} = \binom{n}{n-r}$  you can use the number of  $a$ 's instead.

**Example 6:** Expand  $(a + b)^8$ .

**Solution:**

$$\begin{aligned} (a + b)^8 &= a^8 + \binom{8}{1} a^7 b + \binom{8}{2} a^6 b^2 + \binom{8}{3} a^5 b^3 + \binom{8}{4} a^4 b^4 \\ &\quad + \binom{8}{5} a^3 b^5 + \binom{8}{6} a^2 b^6 + \binom{8}{7} a b^7 + b^8. \\ &= a^8 + 8a^7 b + 28a^6 b^2 + 56a^5 b^3 + 70a^4 b^4 \\ &\quad + 56a^3 b^5 + 28a^2 b^6 + 8ab^7 + b^8. \end{aligned}$$

It's possible to define  $\binom{n}{r}$  even when  $n$  is not a positive integer (but  $r$  must be).

We define  $\binom{n}{r} = \frac{n(n-1)\dots(n-r+1)}{r!}$

**Example 7:**  $\binom{-2}{5} = \frac{-2 \cdot -3 \cdot -4 \cdot -5 \cdot -6}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = -6$ .

$$\binom{\frac{1}{2}}{5} = \frac{\frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot \left(-\frac{5}{2}\right) \cdot \left(-\frac{7}{2}\right)}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{1 \cdot 1 \cdot 3 \cdot 5 \cdot 7}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 2^5} = \frac{7}{256}.$$

It can be shown that the Binomial Theorem is valid even when  $n$  is a fraction, or is negative. In these cases the

Binomial expansion goes on forever and we get, not a polynomial, but an infinite power series.

**Example 8:**

$$\begin{aligned} \frac{1}{1+x} &= (1+x)^{-1} \\ &= 1 + \binom{-1}{1}x + \binom{-1}{2}x^2 + \binom{-1}{3}x^3 + \binom{-1}{4}x^4 + \dots \\ &= 1 + \frac{-1}{1}x + \frac{-1 \cdot -2}{2 \cdot 1}x^2 + \frac{-1 \cdot -2 \cdot -3}{3 \cdot 2 \cdot 1}x^3 \\ &\quad + \frac{-1 \cdot -2 \cdot -3 \cdot -4}{4 \cdot 3 \cdot 2 \cdot 1}x^4 + \dots \\ &= 1 - x + x^2 - x^3 + x^4 - \dots \end{aligned}$$

This is the same power series that we get from the sum to infinity of the GP with first term 1 and common ratio  $-x$ .

We are starting to get out of our depth here, because a proper study of power series involves questions of convergence. Such matters are discussed in my notes *Techniques of Calculus*.

**§5.5. Probability**

The concept of probability, or chance, pervades modern society. We're pleased when we hear that rain is very unlikely, or that a reduction in the mortgage rate is quite probable. Those who gamble understand odds, such as 20 to 1, though a gambler doesn't think of this as representing the likelihood of that horse winning. A committed gambler always believes that his horse is

almost certain to win. For him, the odds of 20 to 1 means that for every dollar invested he'll make \$20 profit.

Probability, as a way of quantifying likelihood, is less well understood. It's a number between 0 and 1 where 0 represents an impossible event and 1 represents an certain event. Often we assign probabilities subjectively. "You have one chance in five of surviving 12 months" is the doctor's way of guessing that the probability of survival after 12 months is  $1/5$  or 0.2.

This is just an estimate, but the probability of certain events can be calculated mathematically. This is based on two principles.

(1) When two or more events are **mutually exclusive**, meaning that no two can occur at the same time, then the probability of one of these occurring is the sum of the individual probabilities.

(2) When two events are considered to be **equally likely**, their probabilities are equal.

**Example 9:** The probability of throwing a 5 or a 6 with a single die (one die – two dice) is  $2/6 = 1/3$ . This assumes that the die is well balanced and so each face is equally likely to come up.

**Example 10:** What is the probability of throwing a total of 6 when two dice are rolled?

**Solution:** You might think that since there are 11 possible totals (2 to 12) the probability of each is  $1/11$ . But they are not all equally likely. There's only one way to get a total of 12, and that's by throwing a double 6. But there are 5 ways of getting a total of 6:

$$1 + 5, 2 + 4, 3 + 3, 4 + 2 \text{ and } 5 + 1.$$

The **sample space** in this case – the set of equally likely events – is the set of ordered pairs

$(x, y)$  where  $x$  is the number shown on the first die and  $y$  is the number shown on the second. It doesn't matter which die is the first and which is the second but, for example a 3 and a 4 is twice as likely as a double 6 because it can arise as  $(3, 4)$  or  $(4, 3)$  whereas a double 6 is just  $(6, 6)$ .

The 36 possibilities in this sample space are equally likely so each has a probability of  $1/36$ . The total of 6 can occur in 5 ways, as  $(1, 5)$ ,  $(2, 4)$ ,  $(3, 3)$ ,  $(4, 2)$  and  $(5, 1)$  and so has a probability  $5/36$ .

So if the sample space of equally likely outcomes has size  $N$  and, of these,  $r$  are 'favourable' to a specific event then we can say that

$$\text{Probability of event} = \frac{\# \text{ favourable}}{\# \text{ possible}} \cdot$$

There's a famous paradox that highlights the fact that our perception of probability is easily mistaken. A TV quiz

show has the contestant faced with three identical boxes. Inside one of these boxes, claims the quizmaster, is a thousand dollars. All you have to do is pick a box!

Now it is clear that, based on this information, your probability of winning the prize is  $1/3$ , or one chance in 3. These are very good odds.

But once you have selected a box, and before you open it, the quizmaster opens one of the other two boxes and shows that it's empty. So the prize has to be in one of two other boxes and so your probability of winning the prize would appear to have risen to  $1/2$ . But the quizmaster urges you to change your mind. He puts a lot of pressure on you to change. Common sense tells you that it doesn't matter whether you change your mind or not, but psychological forces predispose you to resist the temptation to change. Imagine how you'd feel if you gave in and changed only to find out that you would have won if you'd stood firm.

In fact you *should* change. Your probability of winning will be  $2/3$  if you do, not  $1/2$ . If you decide not to change your mind your probability of winning is only  $1/3$ .

If you had chosen the correct box at the outset you will definitely lose by swapping. But this will only happen one third of the time.

If you chose the incorrect box then the prize will definitely be in the other box that is not opened. The quizmaster *knows* where the prize is and will never open the box with the prize. So this means that, by swapping, you'll definitely win if you chose the wrong box to start with, and this will happen two times out of three.

## §5.6. Binomial Probability

With the toss of a single unbiased coin there are two equally likely outcomes, so each has probability  $\frac{1}{2}$ . An analogous situation is the gender of a baby, where the probability of a boy or a girl is each approximately  $\frac{1}{2}$ . (In practice the probability of a baby being a boy is slightly more than a half, but in examples of this type we assume that the probability is exactly a half.)

If two coins are tossed then there are four possible outcomes: HH, HT, TH and TT. Since the outcome of the second toss is independent of what happened with the first, these are all equally likely and so each has probability  $\frac{1}{4}$ .

Where two events are **independent** the probability of both happening is the product of the two probabilities of the individual events.

**For independent events,**  
**Proby(A and B) = Proby(A).Proby(B).**

**Example 11:** A coin is tossed and a die is rolled. What is the probability of getting heads with the coin and a 6 with the die?

**Solution:** The probability is  $\frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$ .



There is a great deal of misunderstanding among the general population concerning independence. A fair coin tossed a large number of times will come up heads about 50% of the time. The more tosses, the closer to 50% the proportion of heads will become. This is known as the **Law of Large Numbers**.

If a coin comes up heads 10 times in a row, there's a strong belief by many people that the next toss is much more likely to be tails, because of the Law of Large Numbers. This ignores the fact that coins have no memory of what has gone on before and the 11<sup>th</sup> toss will be the same as the first. (In actual fact you could argue that 10 heads in a row might be evidence that the coin is biased and therefore a head, on the 11<sup>th</sup> toss may be more likely than tails.)

But even with a completely unbiased coin, 10 heads in a row will come up, on average, once in 1024 ( $= 2^{10}$ ) tosses. If you spend an hour tossing a coin you will be highly likely to witness the phenomena of a run of 10 heads.

In the game of ‘two up’ two coins are tossed and the gambler calls “odds” or “evens”. “Evens” means that both coins come up the same way, while “odds” means they come up differently. If the coins are unbiased the probability for each is  $\frac{1}{2}$ . The same would have happened if just one coin was tossed and the caller chose heads or tails. Why do they bother using two coins?



Two coins are used to allow for coins that may be slightly biased. After all, no coin is perfectly balanced. Suppose that the first coin had a probability  $\frac{1}{2} + p$  of coming up heads, where  $p$  is small, and a probability of  $\frac{1}{2} - p$  of coming up tails. Of course  $p$  could be negative, allowing for the possibility of a coin that was slightly biased towards tails. Suppose that the probabilities for the second coin are  $\frac{1}{2} + q$  and  $\frac{1}{2} - q$ , where  $q$  is small.

	<b>Heads</b>	<b>Tails</b>
<b>First Coin</b>	$\frac{1}{2} + p$	$\frac{1}{2} - p$
<b>Second Coin</b>	$\frac{1}{2} + q$	$\frac{1}{2} - q$

The probability of HH is  $(\frac{1}{2} + p)(\frac{1}{2} + q)$  and the probability of TT is  $(\frac{1}{2} - p)(\frac{1}{2} - q)$ .

So the probability of ‘evens’ is:

$$(\frac{1}{2} + p)(\frac{1}{2} + q) + (\frac{1}{2} - p)(\frac{1}{2} - q)$$

which simplifies to  $\frac{1}{2} + 2pq$ . Now, although this doesn't completely even out the bias, if  $p$  and  $q$  are small then  $2pq$  will be negligible.

For example, if the probability of the first coin coming up heads is 0.51 and the probability of the second coin coming up heads is 0.48 then  $p = 0.01$  and  $q = -0.02$ .

Then  $\text{prob}(\text{Evens}) = 0.5 - 0.0004 = 0.4996$ .

If you toss 8 coins what is the chance of getting 4 heads and 4 tails? For each toss the chance is  $\frac{1}{2}$  of it being a head and  $\frac{1}{2}$  of being a tail (assuming the coins are fair). If we are fussed about the order in which the heads and tails occur then the probability of HTHHTHTT, for example, will be  $(\frac{1}{2})^8 = \frac{1}{256}$ . But there are more ways of getting 4 heads and 4 tails than this particular sequence.

Each of the 4 heads will come from 4 of the 8 tosses.

There are  $\binom{8}{4} = \frac{8 \cdot 7 \cdot 6 \cdot 5}{4 \cdot 3 \cdot 2 \cdot 1} = 70$  such choices and for each of these there'll be one sequence that consists of 4 heads and 4 tails. Each such sequence will have a probability of  $\frac{1}{256}$ , so the required probability is  $\frac{70}{256}$ , which is about 0.273.

Now consider the situation where you carry out an experiment, called a **trial**, and you know the probability of an event. Then the binomial theorem can give the distribution of any number of occurrences of that event in

$n$  trials. In the case of coin tosses the trials are the tosses and the probability of heads is  $\frac{1}{2}$ .

In the following theorem the trials are assumed to be independent. This means that the outcome of each trial doesn't depend on the outcomes of the previous trials. For coin tosses, with an unbiased coin, this would certainly be true. For the sex of births, it's not exactly true since boys or girls tend to run in some families. But it is still close enough to the truth to give good approximations of probabilities.

**Theorem 5:** If an event has a probability of  $p$  in a single trial, and  $q = 1 - p$ , then the probability of getting the event  $r$  times with  $n$  independent trials is  $\binom{n}{r} p^r q^{n-r}$ .

**Proof:** The probability of a particular sequence of outcomes, where the event occurs  $r$  times and not the remaining  $n - r$  times is  $p^r q^{n-r}$ . There are  $\binom{n}{r}$  such sequences with the event occurring  $r$  times so the required probability is  $\binom{n}{r} p^r q^{n-r}$ . 🙌 😊

**Example 12:** Give the approximate probabilities for differing numbers of boys for a family with 10 children (assuming that the sex of the babies is independent for each birth).

**Solution:**

<b># boys</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>proby</b>	$\frac{1}{1024}$	$\frac{10}{1024}$	$\frac{45}{1024}$	$\frac{120}{1024}$	$\frac{210}{1024}$	$\frac{252}{1024}$
<b>approx</b>	0.001	0.010	0.044	0.117	0.205	0.246

<b># boys</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>proby</b>	$\frac{210}{1024}$	$\frac{120}{1024}$	$\frac{45}{1024}$	$\frac{10}{1024}$	$\frac{1}{1024}$
<b>approx</b>	0.205	0.117	0.044	0.010	0.001

**Example 13:** Find the probability of throwing at least two sixes with 4 dice.

**Solution:** We assume that the dice are fair, so that the probability of a six is  $\frac{1}{6}$ .

So  $p = \frac{1}{6}$  and  $q = \frac{5}{6}$ .

The probability of getting exactly 2 sixes is:

$$\binom{4}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 \approx 0.116.$$

The probability of getting exactly 3 sixes is:

$$\binom{4}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right) \approx 0.015.$$

The probability of getting exactly 4 sixes is:

$$\left(\frac{1}{6}\right)^4 \approx 0.001.$$

Adding these, the probability of at least 2 sixes is approximately 0.132..

## EXERCISES FOR CHAPTER 5

**Exercise 1:** On my bookshelves I have the seven novels by the Brontë sisters. I wish to select three of them to take away on holiday. In how many ways can I do this?

**Exercise 2:** Seven athletes run a 100 metres race. The result that is recorded consists of the first place, second place and third place. How many different results are possible?

**Exercise 3:** Seven houses are in a row in a suburban street. Three houses have been burgled. Assuming that the break-ins are random, what is the probability that the three burgled houses will be adjacent.

**Exercise 4:** Expand  $(1 + x)^7$  by the Binomial Theorem.

**Exercise 5:** Expand  $(2a - 3b)^4$  by the Binomial Theorem.

**Exercise 6:** What is the probability of throwing 10 coins and getting exactly 5 heads?

**Exercise 7:** What is the probability of throwing 7 coins and getting exactly 3 heads?

**Exercise 8:** A certain species of dog produces more female pups than males, with the probability of a given

pup being female being  $2/3$ . What is the probability that, in a litter of 8 pups, there are more males than females?

## SOLUTIONS FOR CHAPTER 5

**Exercise 1:** The order of the books I take does not matter, so there are  $\binom{7}{3} = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35$ .

**Exercise 2:** Here the order does matter, so the number of different results  $= 7 \cdot 6 \cdot 5 = 210$ .

**Exercise 3:** There are  $\binom{7}{3} = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35$  ways of selecting three houses. If the choice is random there is a probability of  $\frac{1}{35}$  of each selection. Of these there are 5 choices in which the houses are adjacent. (If the houses are represented by the letters A – G then the adjacent choices are ABC, BCD, CDE, DEF, EFG.) Hence the probability of a random choice selecting three adjacent choices is  $\frac{5}{35} = \frac{1}{7}$ .

**Exercise 4:**  $(1 + x)^7 = 1 + 7x + \binom{7}{2}x^2 + \binom{7}{3}x^3 + \binom{7}{4}x^4$   
 $+ \binom{7}{5}x^5 + 7x^6 + x^7$   
 $= 1 + 7x + 21x^2 + 35x^3 + 35x^4 + 21x^5 + 7x^6 + x^7$ .

**Exercise 5:**  $(2a - 3b)^4$

$$\begin{aligned} &= (2a)^4 - 4(2a)^3(3b) + 6(2a)^2(3b)^2 - 4(2a)(3b)^3 + (3b)^4 \\ &= 16a^4 - 96a^3b + 216a^2b^2 - 216ab^3 + 81b^4. \end{aligned}$$

**Exercise 6:** The probability is  $\frac{\binom{10}{5}}{2^{10}} = \frac{252}{1024} = \frac{63}{256} \approx 0.25$ .

**Exercise 7:** The probability is  $\binom{7}{3} \left(\frac{1}{2}\right)^7 = \frac{35}{256} \approx 0.137$ .

**Exercise 8:** Let  $P(k)$  denote the probability of getting  $k$  males.

Then  $P(k) = \binom{8}{k} \left(\frac{1}{3}\right)^k$ .

The probability of getting more males than females is

$$\begin{aligned} &P(5) + P(6) + P(7) + P(8) \\ &= \binom{8}{5} \left(\frac{1}{3}\right)^5 + \binom{8}{6} \left(\frac{1}{3}\right)^6 + \binom{8}{7} \left(\frac{1}{3}\right)^7 + \binom{8}{8} \left(\frac{1}{3}\right)^8 \\ &= \frac{56}{243} + \frac{28}{729} + \frac{8}{2187} + \frac{1}{6561} \\ &= \frac{1512 + 252 + 24 + 1}{6561} = \frac{1789}{6561} \approx 0.273. \end{aligned}$$