4. PROBABILITY AND THE BINOMIAL THEOREM

§4.1. 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 $26 \times 25 \times 24 \times \ldots \times 3 \times 2 \times 1$. Using a calculator you can determine that this number is just over 4×10^{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 $n(n - 1) \ldots 3.2.1$.

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

§4.2. Numbers of Choices

How many ways are there of choosing 2 things out of 5? If we have 5 letters A, B, C, D, E, 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 doesn't 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:

$n(n-1)(n-2)\ \ldots\ (n-r+1)$ if the order matters and

$$\frac{n(n-1)(n-2)\ \ldots\ (n-r+1)}{r(r-1)(r-2)\ \ldots\ 3.2.1}$$ if the order does not matter.

Proof: 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. (A more formal proof would prove it by induction.) ✋☺

If the choice doesn't matter, each choice of r things can be arranged in $r(r-1)(r-2)$ … 3.2.1 ways and so we must divide by this number. We denote this by $\boldsymbol{r!}$ and call it **factorial *r***.

Example 1: 4! = 4.3.2.1 = 24.

We denote the number of choices of r things from n, where the order matters, as $\boldsymbol{{}^{n}P_{r}}$. Where the order doesn't matter we use the symbol $\boldsymbol{{}^{n}C_{r}}$, or more usually by the symbol $\binom{n}{r}$. We can restate Theorem 1 as follows.

Theorem 1 (again):
The number of ways of choosing r things from n, if order matters, is: ${}^{n}P_{r} = \frac{n!}{r!}$.
The number of ways of choosing r things from n, if order doesn't matter, is:

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

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 1: 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 $\binom{24}{8} = \frac{24.23.22.21.20.19.18.17}{8.\ 7.\ 6.\ 5.\ 4.\ 3.\ 2.\ 1}$ possible choices.

If we cancel as far as possible we get $\frac{3.23.11.19.3.17}{1}$ = 735471.

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

There are $\binom{24}{16} = \frac{24.23.22.21.20.19.18.17.16.15.14.13.}{16.15.14.13.12.11.10.\ 9.\ 8.\ 7.\ 6.\ 5.}$ $\frac{12.11.10.9}{4.\ 3.\ 2.\ 1}$ possible choices.

Now notice that we have the factor 16.15.14.13.12.11.10.9 in both the numerator and denominator. If we cancel these we get $\frac{24.23.22.21.20.19.18.17}{8.\ 7.\ 6.\ 5.\ 4.\ 3.\ 2.\ 1}$ 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
$r_1, r_2, \ldots, r_k$, where $r_1 + r_2 + \ldots + r_k = n$, the number of ways of doing this is: $\frac{n!}{r_1!\, r_2! \ldots r_k!}$. 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) $\binom{n}{r} = \binom{n}{n-r}$;

$$(2)\ \binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1};$$

$$(3)\ \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n.$$

Proof: We've already proved (1).

(2) Suppose we have to choose r things from $n + 1$, say r numbers from $\{0, 1, 2, \ldots, n\}$.
Such a choice might include 0 or exclude 0.
How many exclude 0? In this case we have to choose all r of them from $\{1, 2, \ldots, n\}$ and the number of such choices is $\binom{n}{r}$.
How many include 1? In this case we have the remaining $r - 1$ things to choose from
$\{1, 2, \ldots, n\}$ 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 $\binom{n}{r}$
$+ \binom{n}{r-1}$.

(3) Suppose we have to choose a subset from $\{1, 2, \ldots, n\}$. 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 $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n}$.

But in choosing a subset from $\{1, 2, \ldots, n\}$ 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. ✋☺

§4.3. Pascal's Triangle

There is 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 as follows.

```
                    1
                  1   1
                1       1
              1           1
            1               1
          1                   1
        1                       1
      1                           1
    1                               1
  1                                   1
1                                       1
```

Under each pair of adjacent numbers on each row, write their sum:

$$
\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1
\end{array}
$$

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

§4.4. 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 $(a^2 + 2ab + b^2)(a^2 + 2ab + b^2)$ and finally multiplying the answer by $a^2 + 2ab + b^2$ again. But this would be very laborious.

Imagine expanding $(a + b)(a + b)(a + b)(a + b)(a + b)$. Each term will consist of 5 factors, one chosen from each of these 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 factor. This will give the term *abbab*, which simplifies to a^2b^3.

There will be other terms that simplify to a^2b^3. How many? As many ways as there are of choosing 2 factors, from the 5, where the choice is an 'a'. There are $\binom{5}{2}$ = 10 such choices. So altogether the coefficient of a^2b^3 will be 10.

Now the individual terms in this expansion will all have the form a^rb^{5-r} corresponding to *a* choice of *r* *a*'s and ($n - r$) *b*'s. There will be $\binom{5}{r}$ such terms which can be combined as $\binom{5}{r} a^rb^{n-r}$. Finally, *r* can range from 0 to 5. 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^5 and b^5 will each be 1.

So $(a + b)^5 = a^5 + \binom{5}{1} a^4b + \binom{5}{2} a^3b^2 + \binom{5}{3} a^2b^3 + \binom{5}{4} ab^4 + b^5$. Evaluating the binomial coefficients this becomes $(a + b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$.

Notice that these coefficients can be read off from the 5^{th} row of Pascal's Triangle.

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

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

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 2: Expand $(a + b)^8$.
Solution:

$$(a+b)^8 = a^8 + \binom{8}{1} a^7b + \binom{8}{2} a^6b^2 + \binom{8}{3} a^5b^3 + \binom{8}{4} a^4b^4 + \binom{8}{5} a^3b^5 + \binom{8}{6} a^2b^6 + \binom{8}{7} ab^7 + b^8.$$

$= a^8 + 8a^7b + 28a^6b^2 + 56a^5b^3 + 70a^4b^4 + 56a^3b^5 + 28a^2b^6 + 8ab^7 + b^8$.

§4.5. Binomial Probability

If you toss 8 coins what is the chance of getting 4 heads and 4 tails? For each toss the chance is ½ of it being a head and ½ 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 $\binom{8}{4} = \frac{8.}{4.}\frac{7.}{3.}\frac{6.}{2.}\frac{5.}{1.} = 70$ such choices and for each of these there will 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 ½ .

In the following theorem the trials are assumed to be independent. This means that the outcome of each trial does not depend on the outcomes of the remaining trials. For coin tosses 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's 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 3: 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:

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

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

Example 4: 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 $\binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$ $= 6.\frac{25}{1296} \approx 0.116$.

The probability of getting exactly 3 sixes is $\binom{4}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right) =$ $4.\frac{5}{1296} \approx 0.015$.

The probability of getting exactly 4 sixes is $\left(\frac{1}{6}\right)^4 = \frac{1}{1296}$ ≈ 0.001.

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

EXERCISES FOR CHAPTER 4

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's 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 4

Exercise 1: The order of the books I take does not matter, so there are $\binom{7}{3} = \frac{7.6.5}{3.2.1} = 35$.

Exercise 2: Here the order does matter, so the number of different results = 7.6.5 = 210.

Exercise 3: There are $\binom{7}{3} = \frac{7.6.5}{3.2.1} = 35$ 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: $(2a - 3b)^4 = (2a)^4 - 4(2a)^3(3b) + 6(2a)^2(3b)^2 - 4(2a)(3b)^3 + (3b)^4$

$= 16a^4 - 96a^3b + 216a^2b^2 - 216ab^3 + 81b^4$.

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

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

$P(5) + P(6) + P(7) + P(8) = \binom{8}{5}\left(\frac{1}{3}\right)^5 + \binom{8}{6}\left(\frac{1}{3}\right)^6 + \binom{8}{7}\left(\frac{1}{3}\right)^7 + \binom{8}{8}\left(\frac{1}{3}\right)^8$.

$$= \frac{56}{243} + \frac{28}{729} + \frac{8}{2187} + \frac{1}{6561}$$

$$= \frac{1512 + 252 + 24 + 1}{6561}$$

$$= \frac{1789}{6561} \approx 0.273.$$