2 — TWO OR MORE RANDOM VARIABLES

Many problems in probability theory relate to two or more random variables. You might have equipment which is controlled by two computers and you are interested in knowing the probability of both computers failing at once. An important consideration is whether or not the computers fail *independently.* A hardware failure in one computer is unlikely to provoke a failure in the other but a power cut might cause the simultaneous failure of both computers.

Let's start by considering two dice...

Two Random Variables

When two dice are thrown, you can either think about a single outcome which is one of the 36 possibilities 1,1 to 6,6 or you can think of a separate outcome for each die. It is more usual to take the second view and, in any analysis, use *two* random variables, each of which can take one of the 6 possibilities 1 to 6.

The Multiplication Theorem of Probability

What is the probability of throwing two 6s? If both dice are fair, the answer is $\frac{1}{36}$. This result is an application of the multiplication theorem of probability which, in the present case, says that the overall probability is the product of the individual probabilities. The probability of the first die showing 6 is $\frac{1}{6}$ and the probability of the second die showing 6 is also $\frac{1}{6}$ and the product $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$.

Independence and the Multiplication Theorem — I

The multiplication theorem requires the two throwings to be *independent.* Informally this means that the outcome from one die has no influence on the outcome of the other. It seems most unlikely that one die could affect the other but it is easy to imagine a closely analogous case where independence couldn't be so readily assured.

Consider a two-drum fruit machine where, instead of sporting fruit icons, each drum has six segments numbered 1 to 6. If the machine were well set up the two drums would be independent and behave like two dice but if grit gets into the works independence may be compromised. In an extreme case the two drums might stick together in such a way that they always showed the same number; the probability of obtaining two 6s now would simply be $\frac{1}{6}$.

Independence — Formal Definition

Two events A and B are said to be independent if and only if:

$$\mathrm{P}(B \,|\, A) = \mathrm{P}(B)$$

In simple terms, knowing that event A has occurred makes it neither more likely nor less likely that event B has occurred. In particular, knowing that the first die shows 6 makes it neither more nor less likely that the second die shows 6.

Independence can be an important consideration even when there is just a single random variable. Suppose a card is selected from a pack of cards. The probability of the event 'an ace has been drawn' is $\frac{1}{13}$ and the probability that 'an ace has been drawn *given* that a diamond has been drawn' is also $\frac{1}{13}$. Algebraically:

$$\mathrm{P}(\text{ace} \,\big|\, \text{diamond}) = \mathrm{P}(\text{ace})$$

'Drawing an ace' and 'drawing a diamond' are independent events. By contrast, 'drawing the ace of diamonds' and 'drawing a diamond' are *not* independent events:

$$\mathrm{P}(1\diamondsuit \,\big|\, \text{diamond}) \neq \mathrm{P}(1\diamondsuit) \qquad \text{since} \qquad \frac{1}{13} \neq \frac{1}{52}$$

Knowing that you have a diamond makes it more likely that you have the ace of diamonds.

Independence and the Multiplication Theorem — II

If events A and B are independent then, by definition:

$$\mathrm{P}(B \,\big|\, A) = \mathrm{P}(B)$$

Hence:

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

Noting that set intersection is commutative:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,.\,\mathrm{P}(B) \qquad \text{so} \qquad \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \mathrm{P}(A) \qquad \text{and} \qquad \mathrm{P}(A \,\big|\, B) = \mathrm{P}(A)$$

Unsurprisingly, independence (like set intersection) is commutative.
Note that $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,.\,\mathrm{P}(B)$ is the multiplication theorem. Consider two illustrations:

$$\mathrm{P}(\text{diamond} \cap \text{ace}) = \mathrm{P}(\text{diamond})\,.\,\mathrm{P}(\text{ace}) = \frac{1}{4} \times \frac{1}{13} = \frac{1}{52}$$

and:

$$\mathrm{P}(\text{first die } 6 \cap \text{second die } 6) = \mathrm{P}(\text{first die } 6)\,.\,\mathrm{P}(\text{second die } 6) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

The relationship $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,.\,\mathrm{P}(B)$ will not, of course, be satisfied if A and B are not independent:

$$\mathrm{P}(1\diamondsuit \cap \text{diamond}) \neq \mathrm{P}(1\diamondsuit)\,.\,\mathrm{P}(\text{diamond}) \qquad \text{since} \qquad \frac{1}{52} \neq \frac{1}{52} \times \frac{1}{4}$$

Distinguishability

Again considering two dice, what is the probability of throwing a 1 and a 5? Here there is scope for debate. Do you consider 5 and 1 to be the same as 1 and 5? In probability theory you ask whether or not the two dice are *distinguishable*. If the dice are of different colours or are thrown one into a left-hand bucket and the other into a right-hand bucket, then the dice are distinguishable. If there really is no way of telling which is which (or you don't care) then they are deemed indistinguishable.

If the two dice are distinguishable then clearly the probability of throwing a 1 and a 5 is just the same as throwing two 6s and the answer is $\frac{1}{36}$.

If the dice are indistinguishable then the separate probabilities for 1,5 and 5,1 can be added and the overall probability is $\frac{1}{18}$. Spelt out, the probability is:

$$\frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6}$$

This simple expression uses the multiplication theorem twice and Axiom III (the addition rule) once. The addition rule is applicable since the outcomes 1,5 and 5,1 are exclusive.

A Table for two Random Variables

When there are two random variables it is common to use the names X and Y and the values r and s respectively. One way of tabulating the probabilities is:

$$
\begin{array}{cc|c|c|c|c|c|c|c|cc}
 & & & Y & & s & \rightarrow & & & & \\
 & & 0 & 1 & 2 & 3 & 4 & 5 & 6 & & \\
\hline
 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{0}{6} & \\
\hline
X & 1 & 0 & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{6} & \\
\hline
 & 2 & 0 & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{6} & \\
\hline
r & 3 & 0 & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{6} & \mathrm{P}(\{X=r\}) \\
\hline
\downarrow & 4 & 0 & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{6} & \downarrow \\
\hline
 & 5 & 0 & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{6} & \\
\hline
 & 6 & 0 & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{6} & \\
\hline
 & & \frac{0}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & & \\
 & & & \multicolumn{3}{c}{\mathrm{P}(\{Y=s\})} & \rightarrow & & & &
\end{array}
$$

For convenience, r and s are arranged to start from 0 rather than 1 (see page 1.11) and the probabilities in the first row (when $r = 0$) and first column (when $s = 0$) of the array are all zero. The 36 other entries are all of course $\frac{1}{36}$.

Against the right-hand edge there are seven *marginal sums*. Each value is the sum of the seven probabilities in the associated row. All the rows, except the first, sum to $\frac{1}{6}$. Against the bottom edge there are, likewise, seven marginal sums for the columns.

Extending the Notation $\mathrm{P}(\{X = r\})$

With a single die, there are 6 elementary events (or 7 if the contrived special case of throwing a zero is counted). $\{X = r\}$ corresponds to an elementary event and $\mathrm{P}(\{X = r\})$ is the probability of that event.

With two dice, there are 36 elementary events (or 49 if the contrived special cases are counted). $\{X = r\}$ still corresponds to an event but no longer to an elementary event. An elementary event now is the intersection of two events such as $\{X = 4\}$ and $\{Y = 2\}$ written as $\{X = 4\} \cap \{Y = 2\}$.

Going the other way, the event $\{X = 4\}$ is the union of the seven elementary events $\{X = 4\} \cap \{Y = 0\}$, $\{X = 4\} \cap \{Y = 1\}$, up to $\{X = 4\} \cap \{Y = 6\}$. This may be written:

$$\{X = 4\} = \bigcup_{s=0}^{6} \{\{X = 4\} \cap \{Y = s\}\}$$

The $\mathrm{P}(\{X = r\})$ notation can readily be extended to accommodate the new circumstances:

$\mathrm{P}(\{X = 4\} \cap \{Y = 2\})$	means	The probability of throwing a 4 and a 2.
$\mathrm{P}(\{X = r\} \cap \{Y = s\})$	means	The probability of throwing an r and an s.

By the addition rule, the probability of $\{X = 4\}$ is the sum of the probabilities of the elementary events whose union is $\{X = 4\}$:

$$\mathrm{P}(\{X = 4\}) = \sum_{s=0}^{6} \mathrm{P}(\{X = 4\} \cap \{Y = s\})$$

This is one of the marginal sums in the tabular representation on the previous page.

The notation is rather cumbersome and it is common to use $\mathrm{P}(X = r)$ for $\mathrm{P}(\{X = r\})$ and...

$$\mathrm{P}(X = r, Y = s) \qquad \text{or even} \qquad p_{r,s} \qquad \text{for} \qquad \mathrm{P}(\{X = r\} \cap \{Y = s\})$$

The above summation is then more concisely expressed as:

$$\mathrm{P}(X = 4) = \sum_{s} p_{4,s}$$

Notice that 'summing over s' is taken to mean summing over all valid s starting from zero.

The seven marginal sums at the right-hand edge can be expressed:

$$\mathrm{P}(X = r) = \sum_{s} p_{r,s} \qquad \text{where} \quad r = 0, 1, \ldots, 6$$

The Double-Sigma Notation

The sum of the probabilities of all 36 (or 49) elementary events has to be one and this sum is written:

$$\sum_{r} \sum_{s} p_{r,s} = 1$$

The double-sigma notation is used rather as nested FOR-loops are in a programming language. Thus:

$$\sum_r \sum_s p_{r,s} = \sum_r (p_{r,0} + p_{r,1} + p_{r,2} + p_{r,3} + p_{r,4} + p_{r,5} + p_{r,6})$$

The expression could be expanded further by writing the bracketed item on the right-hand side 7 times over and replacing r by 0 in the first, by 1 in the second, and so on.

Independence of Random Variables — I

In the earlier discussion of independence the term was applied to a pair of *events*. The term can be extended to the idea of a pair of *random variables* being independent.

The two random variables X and Y are said to be independent if and only if:

$$\mathrm{P}(X = r, Y = s) = \mathrm{P}(X = r) \,.\, \mathrm{P}(Y = s) \quad \text{for all } r, s$$

A simple inspection of the probabilities of the elementary events in the table shows this to be true. The probability shown in every cell is the product of the marginal sum at the end of the row and the marginal sum at the bottom of the column. In most cases, the product is $\frac{1}{6} \times \frac{1}{6}$ but for cells in the first row or first column at least one of the terms in the multiplication is zero.

Conditional Probability

Consider just the $Y = 2$ column extracted from the tabular representation of the 49 elementary events:

	$Y = 2$
0	$\frac{0}{36}$
1	$\frac{1}{36}$
2	$\frac{1}{36}$
3	$\frac{1}{36}$
$X = 4$	$\boxed{\frac{1}{36}}$
5	$\frac{1}{36}$
6	$\frac{1}{36}$

The highlighted item shows the probability of the elementary event $\{X = 4\} \cap \{Y = 2\}$. The probability is, of course, $\frac{1}{36}$.

Now consider a different question. What is the probability of this elementary event if we *know* that $Y = 2$? This extra knowledge means that any elementary event has to be in the $Y = 2$ column and the required probability can be expressed:

$$\mathrm{P}(\{X = 4\} \mid \{Y = 2\}) = \frac{\frac{1}{36}}{\frac{0}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}}$$

This is the ratio of the probability of the elementary event of interest to the sum of the probabilities of all the elementary events in the column.

This can be written more generally:

$$\mathrm{P}(\{X=4\} \mid \{Y=2\}) = \frac{\mathrm{P}(\{X=4\} \cap \{Y=2\})}{\sum\limits_r \mathrm{P}(\{X=r\} \cap \{Y=2\})} = \frac{p_{4,2}}{\sum\limits_r p_{r,2}}$$

The summation is down the $Y = 2$ column.

Bayes's Theorem

There is nothing special about 2 and 4 of course and the most general form of the previous expression is:

$$\mathrm{P}(\{X=r\} \mid \{Y=s\}) = \frac{\mathrm{P}(\{X=r\} \cap \{Y=s\})}{\sum\limits_k \mathrm{P}(\{X=k\} \cap \{Y=s\})} = \frac{p_{r,s}}{\sum\limits_k p_{k,s}} \tag{2.1}$$

Note that k is used in the summations to avoid using r in two ways.

This version leads naturally to a consideration of a most important theorem due to Bayes. The problem addressed by Bayes's Theorem is this:

Can you derive $\mathrm{P}(A \mid B)$ from $\mathrm{P}(B \mid A)$?

First recall that if A and B are two events:

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

Noting that the set intersection operator is commutative:

$$\mathrm{P}(A \cap B) = \mathrm{P}(B \mid A) \,.\, \mathrm{P}(A)$$

Given this, (2.1) can be rewritten:

$$\mathrm{P}(\{X=r\} \mid \{Y=s\}) = \frac{\mathrm{P}(\{Y=s\} \mid \{X=r\}) \,.\, \mathrm{P}(\{X=r\})}{\sum\limits_k \mathrm{P}(\{Y=s\} \mid \{X=k\}) \,.\, \mathrm{P}(\{X=k\})}$$

Now let A be the event $\{X = r\}$, let B be the event $\{Y = s\}$ and let A_k be the event $\{X = k\}$ then substitute:

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(B \mid A) \,.\, \mathrm{P}(A)}{\sum\limits_k \mathrm{P}(B \mid A_k) \,.\, \mathrm{P}(A_k)} \tag{2.2}$$

This relationship will be the version of Bayes's Theorem used on this course.

Bayes's Theorem has many applications and a conspicuously successful use has been in speech processing.

Illustration of Bayes's Theorem

Suppose the children in a school are classified in two ways, boys/girls and fair-haired/dark-haired. The following information is available:

> Two-thirds of the children are boys of whom half are fair-haired
> One-third of the children are girls of whom three-quarters are fair-haired

You are asked to determine the probability that a fair-haired child is a boy.

Notice what information is missing. You don't know what proportion of children are fair-haired never mind what proportion of those are boys. To set up a 2×2 table of probabilities would require undertaking a fair amount of arithmetic. It is in cases like this (or more ambitious versions of this) that Bayes's Theorem can be useful.

First consider what information is available:

$$\mathrm{P}(\text{boy}) = \frac{2}{3} \qquad \mathrm{P}(\text{fair} \,|\, \text{boy}) = \frac{1}{2} \qquad\qquad \mathrm{P}(\text{girl}) = \frac{1}{3} \qquad \mathrm{P}(\text{fair} \,|\, \text{girl}) = \frac{3}{4}$$

The information required is $\mathrm{P}(\text{boy} \,|\, \text{fair})$.

These are exactly the circumstances for making use of Bayes's Theorem; you want to derive $\mathrm{P}(A \,|\, B)$ from $\mathrm{P}(B \,|\, A)$.

The next step is to set up two random variables: let X represent boy/girl and Y represent fair/dark. Each random variable can take two values and to make use of the summation in Bayes's Theorem the two values in each case need to be 0 and 1. Map as follows:

$$\text{Let } X = 0 \mapsto \text{boy and let } X = 1 \mapsto \text{girl}$$
$$\text{Let } Y = 0 \mapsto \text{fair and let } Y = 1 \mapsto \text{dark}$$

The supplied information now translates into:

$$\mathrm{P}(X=0) = \frac{2}{3} \quad \mathrm{P}(\{Y=0\} \mid \{X=0\}) = \frac{1}{2} \quad \mathrm{P}(X=1) = \frac{1}{3} \quad \mathrm{P}(\{Y=0\} \mid \{X=1\}) = \frac{3}{4}$$

The information required translates into $\mathrm{P}(\{X=0\} \mid \{Y=0\})$

By Bayes's Theorem:

$$\mathrm{P}(\{X=0\} \mid \{Y=0\}) = \frac{\mathrm{P}(\{Y=0\} \mid \{X=0\})\,.\,\mathrm{P}(\{X=0\})}{\displaystyle\sum_k \mathrm{P}(\{Y=0\} \mid \{X=k\})\,.\,\mathrm{P}(\{X=k\})}$$

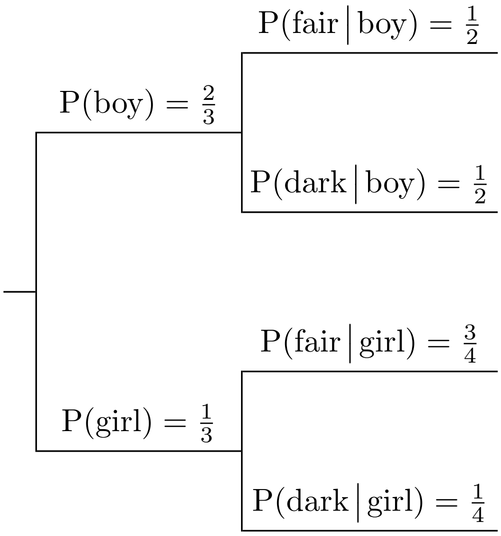
Substituting the known values:

$$\mathrm{P}(\{X=0\} \mid \{Y=0\}) = \frac{\frac{1}{2}\times\frac{2}{3}}{\frac{1}{2}\times\frac{2}{3}+\frac{3}{4}\times\frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{4}} = \frac{4}{7}$$

This is the required answer.

Tree Diagrams

One way to represent the information given in the problem about the children is to use a tree:



$$P(\text{fair} \,|\, \text{boy}) \,.\, P(\text{boy}) = P(\text{fair} \cap \text{boy}) = \tfrac{1}{2} \times \tfrac{2}{3} = \tfrac{1}{3}$$

$$P(\text{dark} \,|\, \text{boy}) \,.\, P(\text{boy}) = P(\text{dark} \cap \text{boy}) = \tfrac{1}{2} \times \tfrac{2}{3} = \tfrac{1}{3}$$

$$P(\text{fair} \,|\, \text{girl}) \,.\, P(\text{girl}) = P(\text{fair} \cap \text{girl}) = \tfrac{3}{4} \times \tfrac{1}{3} = \tfrac{1}{4}$$

$$P(\text{dark} \,|\, \text{girl}) \,.\, P(\text{girl}) = P(\text{dark} \cap \text{girl}) = \tfrac{1}{4} \times \tfrac{1}{3} = \tfrac{1}{12}$$

The tree contains the four known values and the calculations to the right of the tree show four derived values. These are the probabilities of the four elementary events and they can be presented in a 2×2 table:

	fair	dark	
boy	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
girl	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{3}$
	$\frac{7}{12}$	$\frac{5}{12}$	

Once the probabilities of the elementary events are known and entered in the table, the marginal sums can be calculated. The two values at the right-hand edge which correspond to P(boy) and P(girl) were given in the problem and it is worth checking to see that the calculated values are the same. It goes without saying that these should total 1.

The marginal sums at the bottom edge correspond to P(fair) and P(dark). These values also total 1. They are not given in the problem and have not been computed before.

Setting up the table in this way provides an alternative to using Bayes's Theorem to solve the original problem. The two probabilities in the left-hand column at once show that:

$$P(\text{boy} \,|\, \text{fair}) = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{4}} = \frac{\frac{1}{3}}{\frac{7}{12}} = \frac{4}{7}$$

Setting up the table is hard work and it is really very much easier to use Bayes's Theorem.

Independence of Random Variables — II

In the two dice example, the two random variables were independent because:

$$\mathrm{P}(X = r, Y = s) = \mathrm{P}(X = r)\,.\,\mathrm{P}(Y = s) \quad \text{for all } r, s$$

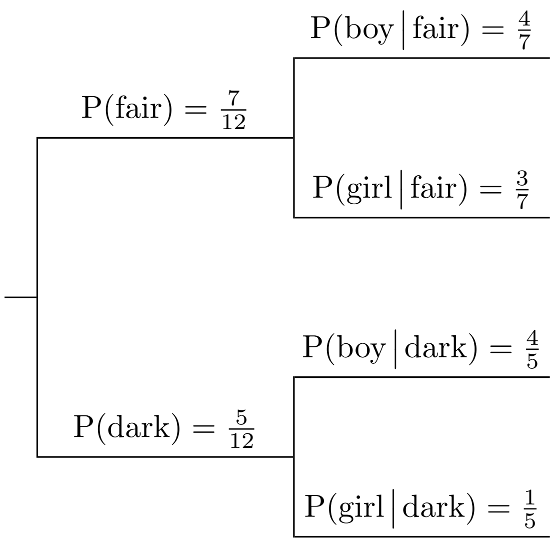
In the problem about the children, simple inspection shows that this identity fails in all four cases:

$$\frac{1}{3} \neq \frac{2}{3} \times \frac{7}{12} \qquad \frac{1}{3} \neq \frac{2}{3} \times \frac{5}{12} \qquad \frac{1}{4} \neq \frac{1}{3} \times \frac{7}{12} \qquad \frac{1}{12} \neq \frac{1}{3} \times \frac{5}{12}$$

Quite clearly the two random variables in this example are not independent.

The Inverse Tree

The first branching in the tree representation on the previous page indicates the division into boys and girls. Now that the table has been constructed, the inverse tree can be constructed. In this the first branching indicates the division into fair and dark:



The value of $\mathrm{P}(\text{boy} \mid \text{fair})$ was determined via Bayes's Theorem earlier and from the table more recently. The other values in the inverse tree are derived from the table.

Event Trees

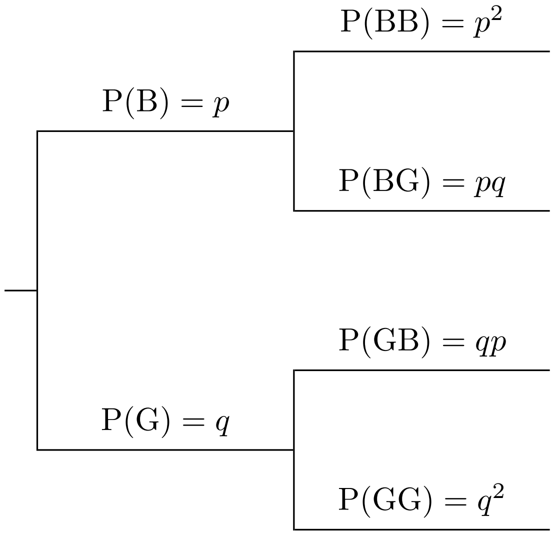
There are many occasions in probability theory when a tree is a good way of presenting information. A second common form of tree is the *event tree*.

Consider a family in which there are two children. In how many ways can the two sexes occur? Trivially, the first child may be a boy or a girl and, in each case, the second child might also be a boy or a girl. This gives rise to four cases (events) which can be represented as BB, BG, GB and GG.

Suppose the probability of a child being a boy is p and the probability of a child being a girl is q where $p + q = 1$. In Britain, $p \approx 0.515$ and $q \approx 0.485$. Assuming the sex of the second child is independent of the sex of the first child, the multiplication theorem can be used and $\mathrm{P(BB)} = p^2$.

The probabilities of the other events are $\mathrm{P(BG)} = pq$, $\mathrm{P(GB)} = qp$ and $\mathrm{P(GG)} = q^2$.

This information can be represented in an event tree:



The two events BG and GB are clearly distinguishable but have equal probabilities given that $pq = qp$.

The sum of the probabilities associated with two children is:

$$p^2 + pq + qp + q^2 = p^2 + 2pq + q^2 = (p + q)^2 = 1$$

Once again, remember that it is good practice to check that the total probability is 1.

The coefficients of the terms in the second expression are 1, 2 and 1 and these are the numbers of ways of having two boys, one of each, and two girls respectively. They are the values in row two of *Pascal's Triangle* and are also known as *Binomial coefficients*.

Combinatorial Numbers — Pascal's Triangle

Here are the first five rows of Pascal's Triangle shown in two ways:

$$\begin{array}{ccccccccc} &&&&1&&&& \\ &&&1&&1&&& \\ &&1&&2&&1&& \\ &1&&3&&3&&1& \\ 1&&4&&6&&4&&1 \end{array} \qquad \begin{array}{c} {}^0C_0 \\ {}^1C_0\ {}^1C_1 \\ {}^2C_0\ {}^2C_1\ {}^2C_2 \\ {}^3C_0\ {}^3C_1\ {}^3C_2\ {}^3C_3 \\ {}^4C_0\ {}^4C_1\ {}^4C_2\ {}^4C_3\ {}^4C_4 \end{array}$$

The rows are numbered from 0 and the values in row two are 1, 2 and 1 as noted in the previous section. In row n there are $n + 1$ values which are numbered 0 to n.

In the version on the right, a given value is identified by ${}^{n}C_{r}$ where n is the row number and r is the position in the row. ${}^{n}C_{r}$ is the number of ways in which there can be r girls in a family of n children.

The symmetry of the values in Pascal's Triangle means that, in general, ${}^{n}C_{r} = {}^{n}C_{n-r}$. For example, ${}^{4}C_{0} = {}^{4}C_{4}$ and ${}^{4}C_{1} = {}^{4}C_{3}$. In words: the number of ways there can be $n - r$ girls in n children, ${}^{n}C_{n-r}$, is the same as the number of ways there can be r boys in n children which is exactly the same as the number of ways there can be r girls in n children, ${}^{n}C_{r}$.

As already noted, with two children there are 2 ways in which one may be a girl, BG and GB, and ${}^{2}C_{1} = 2$. Pascal's Triangle shows that with four children there are 6 ways in which two may be girls, ${}^{4}C_{2} = 6$. The values in Pascal's triangle are examples of *combinatorial numbers* and the C stands for *Combinations*.

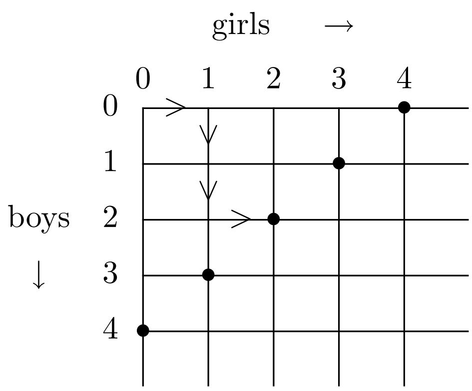
Pascal's Theorem states that:

$$ {}^{n+1}C_{r+1} = {}^{n}C_{r} + {}^{n}C_{r+1} $$

Informally, each value is the sum of the two nearest values in the row above. This doesn't apply to the edge values which are all 1.

Grid or Lattice Representation

Another useful form of representation when analysing problems in probability theory is to use a *grid* or *lattice*. Here is an example which can be used to determine combinations of boys and girls:



The five blobs constitute the sample space of elementary events when there are four children and boy-girl order is not relevant. The five blobs represent 4 boys and 0 girls, 3 boys and 1 girl and so on.

This representation can be used to determine combination values. For example, suppose you wish to determine ${}^{4}C_{2}$, the number of ways that there can be two girls in four children. You simply count the number of ways of getting from the top left-hand corner to the appropriate blob using only rightward and downward steps.

In the present case, the middle blob of the five represents two girls in four children and one path from the top left-hand corner to this blob is marked. This one goes right-down-down-right. Check that you can find the other five routes.

The Binomial Theorem

The Binomial Theorem is essentially the expansion of $(x+y)^n$:

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

The item $\binom{n}{r}$ (pronounced n choose r) is a Binomial coefficient and has the same value as ${}^nC_r$, a combination. The two forms can be expressed using factorials:

$$\binom{n}{r} = {}^nC_r = \frac{n!}{r!\,(n-r)!}$$

Glossary

The following technical terms have been introduced:

independent	Pascal's Triangle	combination
distinguishable	Binomial coefficient	grid
marginal sum	combinatorial number	lattice
event tree		

Exercises — II

Again work in fractions for all the following problems which have a numerical content. Treat, for example, 99% as $\frac{99}{100}$ and not as 0.99.

1. A bag contains one ball which, equiprobably, may be black or white. A white ball is added to the bag which is then shaken. One ball is retrieved at random and found to be white. What is the probability that the other ball is white?

2. A shopkeeper obtains light bulbs from three suppliers, A_1, A_2 and A_3 and has noted the following information about the bulbs:

   - 30% come from supplier A_1 and, of these, 1% don't work.
   - 45% come from supplier A_2 and, of these, 1% don't work.
   - 25% come from supplier A_3 and, of these, 2% don't work.

   For a bulb selected at random, consider two random variables S (the Supplier which may be A_1, A_2 or A_3) and D (the Dudness which may be G or B for Good or Bad). Note that the probability of a randomly selected light bulb from Supplier A_1 being Bad can be expressed as $\mathrm{P}(D = B \mid S = A_1) = \frac{1}{100}$.

   Tasks:

   (a) Express the information provided in Tree Representation. From the root of the tree there should be three branches (one for each supplier). Label these branches with probabilities, $\frac{30}{100}$, $\frac{45}{100}$ and $\frac{25}{100}$ respectively. From each node (there is one at the end of each branch) draw two more branches; these should be labelled with the appropriate conditional probabilities of Dudness.

(*b*) From the tree representation, derive a Tabular Representation. The table should have three rows (one for each Supplier) and two columns (headed G and B for Dudness). Each of the six cells in the table should show the probability of a randomly selected bulb having come from the relevant Supplier and having the relevant Dudness.

(*c*) Mark in the row totals. These should be $\frac{30}{100}$, $\frac{45}{100}$ and $\frac{25}{100}$.

(*d*) Mark in the column totals. Note that the total of the B column shows the overall probability of a randomly selected bulb being Bad.

(*e*) Given that a randomly selected bulb is Bad, what is the probability that it came from Supplier A_3?

(*f*) From the tabular representation derive the alternative tree representation. From the root of this tree there should be two branches (one for Good and one for Bad) labelled with the two appropriate probabilities (the Bad branch should be labelled with the probability determined in Task *d*). From each node there should be three branches (one for each Supplier) each labelled with the appropriate conditional probability.

(*g*) Redraw the tabular representation, but replace the numerical values in the six cells by expressions of the form $\mathrm{P}(B \,|\, A_1) \times \mathrm{P}(A_1)$ (the first part of this being an abbreviation of the expression given in the preamble).

(*h*) Use the revised tabular representation to illustrate Bayes's Theorem.

3. One person in 1000 is known to suffer from Nerd's Syndrome (a pathological inability to resist playing computer games). A standard test is such that 99% of those who suffer from Nerd's Syndrome show a positive result. The same test also (falsely) shows positive with 2% of *non*-sufferers.

   A person selected at random is tested and found positive. What is the probability that this person actually suffers from Nerd's Syndrome?

4. A coin is tossed until the first time the same result appears twice in succession. To every possible outcome requiring n tosses attribute equal probability. Describe the sample space. Find the probability of the following events: (*a*) the experiment ends before the sixth toss, (*b*) an even number of tosses is required.

5. What is the probability that two throws with three dice show the same configuration if (*a*) the dice are distinguishable and (*b*) they are not.

6. From first principles, prove Pascal's Theorem that ${}^{n+1}C_{r+1} = {}^{n}C_r + {}^{n}C_{r+1}$. Do not make use of the representation which uses factorials. Instead reason as follows:

   Assume that ${}^{n+1}C_{r+1}$ really is the number of ways of having $r+1$ girls in a family of $n+1$ children and consider, separately, the number of these ways where the last child is a girl and the number of these ways where the last child is a boy.

7. [From Part IA of the Natural Sciences Tripos, 1996 (Elementary Mathematics for Biologists)] Alice is observing the White Knight and the Red Knight, who are fighting for the privilege of taking her prisoner. They have agreed to observe the Rules of Battle, which are as follows:

   In each round, the White Knight hits the Red Knight with probability $\frac{2}{10}$, and the Red Knight hits the White Knight with probability $\frac{3}{10}$. It is not possible for both to score a hit in the same round. If the Red Knight is hit, he falls off his horse with probability $\frac{6}{10}$; if the White Knight is hit, he falls off his horse with probability $\frac{8}{10}$.

   A Knight who scores a hit never falls off his horse but if the Knights miss one another then either may fall off his horse from surprise; they do so independently, the Red Knight with probability $\frac{4}{10}$, and the White Knight with probability $\frac{6}{10}$.

   If they both fall off in the same round the battle ends; they shake hands and walk away (forgetting about Alice).

   In all other cases the Knights proceed to a further round.

   (*a*) What is the probability that the Red Knight falls off his horse in the first round?

   (*b*) What is the probability that both fall off in the first round?

   (*c*) What is the probability that the battle ends after exactly three rounds?

8. Four married couples go to a wife-swapping party. The names of the wives are thrown into a hat, shuffled, and then distributed randomly one name to each husband. Each man goes home with the woman allocated to him. What is the probability that no husband goes home with his own wife?

9. Suppose n couples go to a wife-swapping party organised as in the previous question. What now is the probability that no husband goes home with his own wife? It is suggested that you proceed as follows...

   Suppose A_i is the event 'Man$_i$ goes home with his own wife'.

   Then $A_1 \cup A_2 \ldots \cup A_n$ is the event 'At least one man goes home with his own wife'.

   The required probability (that no man goes home with his own wife) can then be expressed as:

$$\text{Probability} = 1 - \mathrm{P}(A_1 \cup A_2 \ldots \cup A_n)$$

   This requires recourse to the inclusion-exclusion theorem and, in the four-couple case of the previous question, one may exploit the solution to Exercise I, question 11, to obtain:

$$\begin{aligned} \text{Probability} = 1 & & \text{1 term, value } 1 \\ & -[\mathrm{P}(A_1) + \mathrm{P}(A_2) + \ldots] & \text{4 terms, each } \tfrac{1}{4} \\ & +[\mathrm{P}(A_1 \cap A_2) + \ldots] & \text{6 terms, each } \tfrac{1}{4} \times \tfrac{1}{3} \\ & -[\mathrm{P}(A_1 \cap A_2 \cap A_3) + \ldots] & \text{4 terms, each } \tfrac{1}{4} \times \tfrac{1}{3} \times \tfrac{1}{2} \\ & +[\mathrm{P}(A_1 \cap A_2 \cap A_3 \cap A_4)] & \text{1 term, value } \tfrac{1}{4} \times \tfrac{1}{3} \times \tfrac{1}{2} \times \tfrac{1}{1} \end{aligned}$$

   Evaluate this expression and confirm that the result agrees with your answer to the four-couple case. It should now be easy to write down the general expression for the n-couple case which you should then prove by induction.