Teach Yourself
Basic Probability

This document is intended as a simple introduction to the subject for those who have not met probability theory as part of their previous maths studies. Probability theory is one of those mathematical topics which is best learnt from seeing and performing a large number of examples. Each new topic covered here is illustrated by a number of worked examples. If the material is new to you, these should be followed carefully and then the corresponding basic exercises attempted. Worked solutions for these exercises are included in a later section (Appendix A).

Further worked solutions are provided in Appendix B (which can also be used as exercises).

The final section of the document (Appendix C) is an examples sheet which should be prepared for a supervision in the normal manner.

Those who have studied probability before may find it sufficient to move straight to the examples sheet.

N.B. $x!$, $_nC_r$, $_nP_r$, *mean* and *sample standard deviation* appear as buttons on standard calculators.

Contents

Recommended books

'Statistics for Advanced Level' by Jane Miller,
            published by C.U.P. (Second Edition 1989)
Probability Theory   pp 54 – 76
Permutations and Combinations   pp 83 – 88

'Mathematics - The Core Course for A-level' by L. Bostock and S. Chandler,
            published by Stanley Thornes (Publishers) Ltd. (1981)
Permutations and Combinations   Chapter 14

'Modern Engineering Mathematics' by Glyn James
            published by Addison-Wesley (Fourth Edition 2008)
Probability Theory   pp 973 – 1035

'Advanced Engineering Mathematics' by Erwin Kreyszig,
            published by John Wiley (Eighth Edition 1999)
Probability and Statistics   pp 1049 – 1155

PROBABILITY OF AN EVENT

When tossing an unbiased coin, it is fairly obvious what we mean by 'the probability of getting heads is 50%'. This is a statement that, if we were to toss a coin a large number of times, then the proportion of tosses which result in heads will approach one half. This intuitive notion of probability as the relative frequency of something is a good starting point.

Example 1

Toss a coin three times, what is the probability of at least two heads ?

Answer

There are 8 possible outcomes which, if the coin is unbiased, should all be equally likely:-

$$\boxed{\text{H H H}} \quad \boxed{\text{H H T}} \quad \boxed{\text{H T H}} \quad \text{H T T}$$

$$\boxed{\text{T H H}} \quad \text{T H T} \quad \text{T T H} \quad \text{T T T}$$

Two or more heads result from the 4 outcomes which are ringed.

The probability of two or more heads is, therefore:

$$\text{Probability} = \frac{4}{8} = \frac{1}{2}$$

We solved this problem by first enumerating the set of possible outcomes, known as the **Sample Space**, and then by deciding which of these outcomes satisfied the criterion of containing 'two or more heads'. A particular subset of outcomes such as 'two or more heads' is conventionally known as an **event** and denoted A, with the probability of the event A denoted **P(A)**. In this case $P(A) = 0.5$ or 50%.

Example 2

A bag containing lettered Scrabble pieces has the following letter distribution

A	B	C	D	E	F	G	H	I	J	K	L	M
9	2	2	4	12	2	3	2	9	1	1	4	2
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
6	8	2	1	6	4	6	4	2	2	1	2	1

The first letter is chosen at random from the bag. Find the probability that it is: (i) an E; (ii) in the first half of the alphabet; (iii) in the second half of the alphabet;
(iv) a vowel; (v) a consonant; (vi) the only one of its kind.

Answer

The total number of pieces in the bag is (by simple addition) 98.

(i) 12 are E's, giving $P(E) = \dfrac{12}{98}$

(ii) By addition, the total number of pieces with A to M is 53, hence $P(\text{1st half}) = \dfrac{53}{98}$

(iii) If there are 53 pieces corresponding to the first half of the alphabet, there must be 98 - 53 = 45 corresponding to the second half. Hence $P(\text{2nd half}) = \dfrac{45}{98}$

(iv) Number of vowels = 9 + 12 + 9 + 8 + 4 = 42 $\Rightarrow P(\text{vowel}) = \frac{42}{98}$

(v) Number of consonants = 98 - 42 = 56 $\Rightarrow P(\text{consonant}) = \dfrac{56}{98}$

(v) Letter must be one of J, K, Q, X or Z $\Rightarrow P(\text{only one of kind}) = \dfrac{5}{98}$

In general, then, probability is relative frequency:

Possible outcomes = N \qquad Number of outcomes for which A happens = $n(A)$

$$\text{Define } P(A) = \frac{n(A)}{N}$$

A number of intuitively reasonable properties follow immediately from this definition. The first is that

$$0 \le P(A) \le 1$$

and as expected of a measure of how likely an event is to happen:

$P(A) = 0$ implies A *never* happens and
$P(A) = 1$ implies A *always* happens

In solving example 2, part (ii), we reduced the labour of counting by making use of the fact that a letter was either in the first half of the alphabet or the second. This line of reasoning was used again in part (iv) when deciding that a letter was either a vowel or a consonant. In formal terms this can be stated as:-

If 'not A' is denoted by $\bar{A}$, the complement of A in the sample space, then between them A and 'not A' cover all cases

$$P(\bar{A}) + P(A) = 1$$

Now try the following exercises. You should, in general, try to complete the exercises as they appear in the text, before moving on to the next section. Written solutions for all exercises appear in Appendix A (page 15).

Exercise 1
One card is drawn from a standard pack of 52 playing cards. What is the probability of
 (a) picking a red card   (b) picking a king   (c) picking a diamond ?

ADDITION OF PROBABILITIES

A further glance at Example 2 indicates that, in deciding on the probability of a vowel, we reasoned that

vowel = A or E or I or O or U

and thus

$$n(\text{vowels}) = n(A) + n(E) + n(I) + n(O) + n(U)$$
$$P(\text{Vowel}) = P(A) + P(E) + P(I) + P(O) + P(U)$$

so probabilities of different events simply add. Or do they ?

Example 3
What is the probability of drawing an ace or a spade from a well-shuffled pack of cards ?

Answer
There are 4 aces so $P(\text{ace}) = \dfrac{4}{52}$.   There are 13 spades: $P(spade) = \dfrac{13}{52}$.
It follows that the probability of an ace or a spade $= \dfrac{4+13}{52} = \dfrac{17}{52}$.

This is, in fact, incorrect . The problem is that the ace of spades has been counted twice, once as an ace and once as a spade.

The correct answer is $\dfrac{4+13-1}{52} = \dfrac{16}{52} = \dfrac{4}{13}$

The First Great Pitfall

The temptation to add probabilities for different events without checking for double counting is a common error. A number of pictorial ways of guarding against making this mistake are useful. 'Tree' or 'branching' diagrams are probably the foremost of these.

Example 4
Around 0.8% of men are blue-green colour-blind (the figure is slightly different for women) and roughly 1 in 5 men is left-handed. Assuming these characteristics are inherited independently, calculate the probability that a man chosen at random will:

       (i) be both colour-blind and left-handed
       (ii) be colour-blind and not left-handed
       (iii) be colour-blind or left-handed
       (iv) be neither colour-blind nor left-handed.

Answer

The tree diagram indicates the way the male population is divided. First of all into left-handed (L) and right-handed (R) groups which contain 20% and 80% respectively. The left-handed group then divides into colour-blind (CB) and not ($\bar{C}B$) comprising 0.8% and 99.2% of the left-handed group. The right-handed group divides in a similar manner.
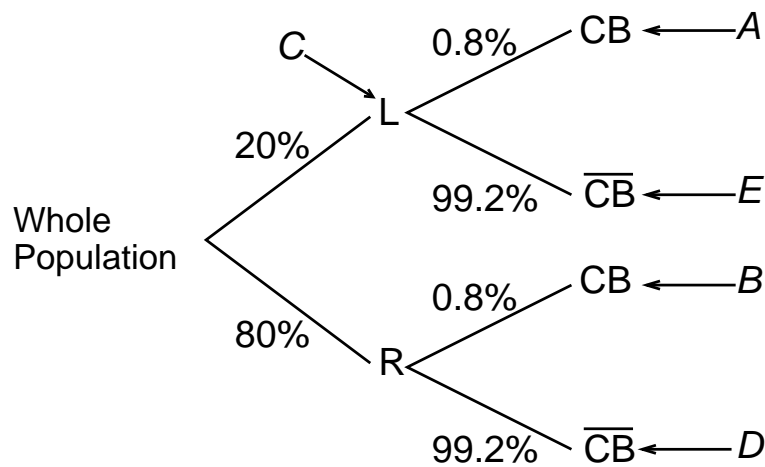
(i) the percentage of those colour-blind and left-handed are those ending up at A. There are $0.008 \times 0.2 = 0.0016$ of the population in this category. i.e Required probability $= 0.0016$.

(ii) This is group B. $P(B) = 0.008 \times 0.8 = 0.0064$.

(iii) This is group C plus group B. Required probability $= 0.0064 + 0.2$

(N.B. there are no individuals in both B & C - all branches of a tree should be exclusive)

(iv) This is group D. $P(D) = 0.992 \times 0.8 = 0.7936$.

Note that we could deduce the probability of group E now, as $1 - P(A) - P(B) - P(D)$.

Since events are sets of outcomes, Venn diagrams are also occasionally useful. This is particularly the case in deriving general formulae:
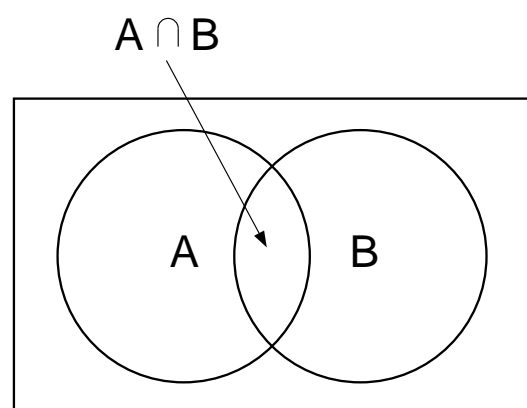
If A and B represent two events (i.e. two sets of outcomes), when considering outcomes which are either in A or in B (or both), that is $A \cup B$, if we add the outcomes in A to those in B, we have counted those in both (i.e. $A \cap B$) twice. The general rule for addition of probabilities is:-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
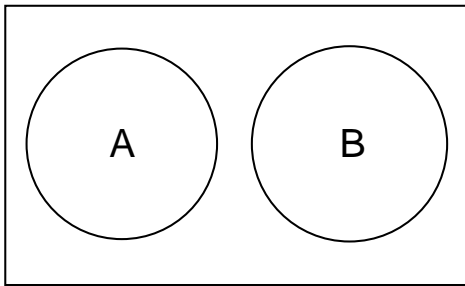
Probabilities can safely be added for events A and B that cannot both happen, i..e. are **mutually exclusive**

Exclusive events $P(A \cup B) = P(A) + P(B)$

$A \cup B = \emptyset \quad P(A \cap B) = 0$

Exercise 3

100 cars are entered for a road-worthiness test which is in two parts, mechanical and electrical. A car passes only if it passes both parts. Half the cars fail the electrical test and 62 pass the mechanical. 15 pass the electrical but fail the mechanical test. Find the probability that a car chosen at random

(i) passes overall   (ii) fails on one test only

(iii) given that it has failed, failed the mechanical test only .

MULTIPLICATION OF PROBABILITIES

In example 4 above, to calculate the proportion of the male population who were both left handed and colour-blind (those in group A), we reasoned that 0.2 of the population were left-handed and of this group of left-handers 0.008 were blue-green colour-blind. In terms of probabilities, this translates to:-

proportion left-handed and colour-blind = proportion left-handed $\times$ proportion colour-blind

$$\Rightarrow P(L \cap CB) = P(L)P(CB)$$

and, in an identical manner, for those right-handed and colour-blind (in group B)

$$P(R \cap CB) = P(R)P(CB)$$

It seems clear then that to calculate the probability of two events both happening we simply multiply the separate probabilities that each happens.

There was, however, a crucial further piece of information used in this reasoning. This was that *"these characteristics are inherited independently"*. We were thus able to say that the proportion of left-handers who are colour-blind is the same (0.008) as in the total male population (as it also is for right-handers). If the two genetic traits, handedness and colour-blindness, are *not* inherited independently, then this reasoning is flawed and we can progress no further without some knowledge of the dependence.

For **independent** events A and B though we are safe in saying

$$P(A \cap B) = P(A)P(B)$$

Exercise 4

Explain the flaw in this argument and rewrite it as a valid statement. *The probability of throwing a 6 on a fair die = 1/6. Therefore the probability of throwing at least one six in six tosses of the die is 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1 so it is a certainty.*

The Second Great Pitfall

Multiplying probabilities of events is such an intuitively natural, and often correct, thing to do that it is easy to lose sight of the fact that not all events are independent. This is the second commonest mistake in calculating probabilities.

Example 5

Suppose 0.01% of the population have a rare genetic disorder. A test is designed which can detect the condition, if it is present, with 95% probability. The test gives the occasional false positive reading amongst people without the disorder, with a probability of 0.05%.

Find the probability that a person has the disorder given that the result of the test is positive.

Answer

Let D = has the disorder (with $\bar{D}$ representing does not have it) and let R = positive result (so that $\bar{R}$ is a negative result). The statistics for the percentage of the population with the disorder imply $P(D) = 0.0001$ and $P(\bar{D}) = 0.9999$. If we use a tree diagram, the population divides as shown. We can then represent the results of the test applied to each group separately by the second level of branches.
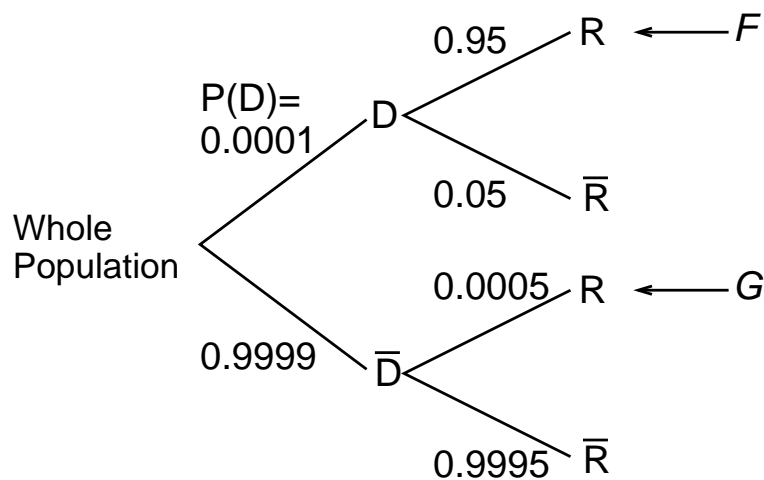
The proportion of those with the disorder (D) who test positive has been marked as 0.95, meaning that 0.05 of this group, although having the disorder, do not test positive. Those free of the disorder test positive and negative with probability 0.0005 and 1 - 0.0005 = 0.9995 respectively.

If a positive result is obtained, then on the diagram it is in branch F or G. Now, if the whole population were tested, the proportion of tests ending in group F would be $0.0001 \times 0.95 = 0.000095$ and that in group G $0.9999 \times 0.0005 = 0.00049995$.

The required probability, which is the proportion of all those testing positive who actually have the disorder, is thus

$$\frac{n(F)}{n(F) + n(G)} = \frac{0.95 \times 0.0001}{0.95 \times 0.0001 + 0.0005 \times 0.9999} = 0.16$$

(Despite the relatively high accuracy and small rate of false positives, this test is useless since 5 out of every 6 who test positive are, in fact, clear of the disorder.)

CONDITIONAL PROBABILITY

Example 5 is relatively straightforward once we realise that we have to distinguish between the probability of testing positive amongst those who do and those who do not have the disease. The technical term for doing this is *conditional probability*. This is defined as the "probability of event A given that event B is *known to have happened*", denoted **P**(**A**|**B**) and usually abbreviated to the "probability of A given B".

In terms of conditional probabilities, the various pieces of information given in example 5 can be written as follows: $P(D) = 0.0001$; $P(R|D) = 0.95$ & $P(R|\bar{D}) = 0.0005$, and the purpose of the exercise is to find $P(D|R)$.

If we label the branching diagram with the various conditional probabilities, with, at each stage, a branch being labelled with the probability of taking that branch, then we obtain
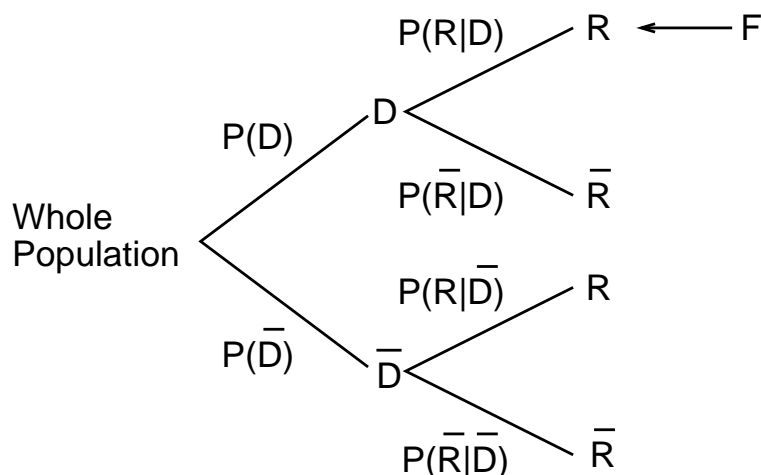
The probability of ending up in group F is precisely $P(R \cap D)$ and it is clear from the branching diagram that conditional probabilities multiply

$$P(R \cap D) = P(R|D)P(D)$$

i.e.

$$P(R|D) = \frac{P(R \cap D)}{P(D)}$$

A number of valid chains of reasoning also become apparent. For example, from the branching diagram it is easy to see that

$$P(R) = P(R|D)P(D) + P(R|\bar{D})P(\bar{D})$$
$$P(R|D) + P(\bar{R}|D) = 1 \text{ etc.}$$

If you see enough examples these rules will begin to seem perfectly natural. For the moment, I would not bother committing them (other than the ones boxed) to memory but would draw a quick sketch of the branching diagram where necessary.

Exercise 5
Four suppliers provide 10%, 20%, 30% and 40% of the bolts sold by a hardware shop and the rate of defects in their products are 1%, 1.5%, 2% and 3% respectively. Calculate the probability of a given defective bolt coming from supplier 1.

INDEPENDENCE

We noted above that probabilities for independent events A and B could be multiplied to find the probability of both happening. This is not, however, a working definition of independence. A more natural definition of independence is in terms of conditional probabilities. Two events A and B are independent if the probability that A happens is not influenced by whether or not B happens,

$$P(A|B) = P(A|\bar{B}).$$

There is, moreover, no need to distinguish either of these conditional probabilities from P(A).

Remember, if A and B are independent then

$$P(A \cap B) = P(A)P(B)$$

and *not otherwise*.

Example 6
A biased coin falls with heads showing with probability p. The coin is tossed three times. Find the probability of obtaining:-

(i) three heads   (ii) two heads and one tail.

Answer

(i) It is reasonable to assume that the tosses are independent and the probability of getting a head on any one is p. Obtaining a head on the three tosses are three independent events and we multiply the three probabilities, giving p^3.

(ii) This can occur as HHT HTH or THH . The probability of throwing tails is $1 - p$, and so the probability of throwing each of these combinations is $p^2(1 - p)$. These three sequences can not occur together (i.e. they are exclusive) and hence required probability is

$$3p^2(1 - p)$$

Exercise 6

The two events A and B have $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$, $P(A \cup B) = \frac{2}{3}$.

Are the events A and B (i) independent (ii) mutually exclusive ?

PERMUTATIONS

The calculation of a probability often reduces to working out the number of ways it is possible to produce a certain event. In earlier examples, e.g. Example 1, this was accomplished by simply listing all the possible outcomes and counting. This is not feasible when the number of outcomes becomes large and some other general methods are needed.

Example 7

If three cards are dealt, what is the probability that the first is the ace of spades, the second the ace of hearts, and the third the ace of clubs?

Answer

Of all the possible deals, only one results in this hand. The required probability is thus

$$P(\text{this hand}) = \frac{1}{\text{Number of deals}}$$

and it only remains for us to calculate the total number of possible deals.

To do this, we note that the number of different possibilities for the first card = 52. For *each choice of first card*, the number of different possibilities for the second = 51 and then, for *each pair of choices for the first two cards*, the number of different outcomes for the third = 50.

The total number of distinct deals = $52 \times 51 \times 50$

and probability of the required deal = $\dfrac{1}{52 \times 51 \times 50} = 7.54 \times 10^{-6}$.

9

It is quite important, I think, to come to grips with this type of reasoning and it is worth spending some time thinking about this example, following through the reasoning making concrete the choice of deal. Suppose the first card dealt is the 6♦, there are 51 cards left. Suppose the second is the 3♡, there are now 50 cards left and hence 50 possible deals in which the first and second cards are 6♦ and 3♡. There is, however, nothing special about the 6♦ and the 3♡ and there will be 50 deals for any particular choice of the first two.

There are in example 7 the seeds of a general rule.

Example 8

n runners enter a cross-country race. If prizes are awarded to the first r to finish, in how many ways can the prizes be distributed ?

Answer
Number of possibilities for the first prize $= n$
For each winner, the number of choices for second $= n - 1$ etc.
Finally, for each set of 1 to $r - 1$'th prize winners there are $n - r + 1$ choices for the r'th.
The total number of ways the prizes can be distributed is thus

$$n(n-1)(n-2)\ldots(n-r+1)$$

This repeated product of integers one smaller than the last occurs regularly in this business and it is much more convenient to work in *factorial* notation. With "n factorial", denoted $n!$, defined as

$$n! = n(n-1)(n-2)\ldots 3 \times 2 \times 1$$

the product above can be simplified to

$$n(n-1)(n-2)\ldots(n-r+1) = n(n-1)(n-2)\ldots(n-r+1)\frac{(n-r)(n-r-1)\ldots 1}{(n-r)(n-r-1)\ldots 1}$$

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

i.e. The number of ways of distributing r prizes to n runners is $\dfrac{n!}{(n-r)!}$

Exercise 7
Write in factorial notation

$$\text{(i)}\frac{8 \times 7 \times 6}{5 \times 4 \times 3} \qquad \text{(ii)}\frac{15 \times 16}{4 \times 3 \times 2} \qquad \text{(iii)}\frac{(n-1)n(n+1)}{4 \times 3 \times 2}$$

and show that

$$9 \times 7 \times 5 \times 3 \times 1 = \frac{9!}{2^4 4!}$$

The Third Great Pitfall

Example 8 exemplifies how to choose the *r* prize winners from the n runners and is thus the number of ways of choosing *r* objects from a group of *n*. Or is it ?

The answer to this question is *sometimes*. The prizes for a race are *all different* with the first prize of greater value than the second which is greater than the third, etc. The *order* in which the *r* prize winners are chosen from the n runners clearly matters, and an order of merit consisting of the *same set of r* but in a *different order* would be regarded as a different way of distributing the prizes. This is in contrast, for example, to the number of ways a shortlist of *r* job applicants can be chosen from a field of *n*. All orderings of the shortlist would, in advance of interview, be regarded as the same. It is extremely easy to forget whether or not the order matters when calculating "the number of ways of choosing *r* objects from *n*".

When the order matters, we usually refer to the "number of ways of **permuting** *r* objects from *n*", and each selection is referred to as a **permutation** of *r* objects from *n*.

The number of ways of permuting (i.e. choosing such that order matters) *r* objects from a total of *n* is denoted $_n\mathbf{P_r}$ or $^n\mathbf{P_r}$ and

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

When order doesn't matter, we refer to the "number of ways of **combining** *r* objects from *n*", and each selection is referred to as a **combination** of *r* objects from *n*.

We will see, from looking at a number of examples, that combinations are used more often than permutations and the term "choosing" usually refers to "combining".

COMBINATIONS

Example 9
What is the probability of being dealt a three card hand which consists of the aces of hearts, spades and clubs ?

Answer
This is essentially the same as example 7, but this time the order in which the aces are dealt *does not matter*.
The separate deals which result in the required hand are

HSC   HCS   SHC   SCH   CHS   CSH

i.e. There are 6 which result in this hand. The probability of getting this hand

$$\frac{6}{52 \times 51 \times 50} = 4.52 \times 10^{-5}.$$

Rather than listing all the possible deals which lead to the required hand, a more instructive way of calculating the number is to note that
Number of choices for the first ace (from among HSC) = 3
For each choice of first ace, number of choices for the second = 2
and then 1 for the third

Giving $3 \times 2 \times 1 = 3! = 6$ ways of ordering the three aces.

(The number of ways of arranging 3 objects amongst themselves in which the order matters is, in fact, $_3P_3$. This is consistent with the definition

$$_3P_3 = \frac{3!}{(3-3)!} = \frac{3!}{0!}$$

provided we interpret 0! as 1 (N.B. *not* 0)

Example 10
Find the number of ways of combining r objects from n.

Answer
The easiest way of doing this is to generalise the argument used in example 9.
r objects can be arranged amongst themselves (i.e. ordered) in $r!$ ways (there are r choices for the first , $r-1$ for the second, etc)
Each separate *combination* of r objects thus corresponds to $r!$ different permutations.
Hence the number of different ways of combining r articles from n, denoted $_n\mathbf{C_r}$ or $^n\mathbf{C_r}$, is given by

$$_nC_r = \frac{_nP_r}{r!} = \frac{n!}{(n-r)!r!}$$

$_nC_r$ is also denoted $\binom{n}{r}$.

Returning to example 9, we could have argued that:

Total number of different hands (ignoring the order in which the cards are dealt) $= _{52}C_3$.

The probability of getting any particular one is thus $\dfrac{1}{_{52}C_3}$.

Example 11
What is the probability of being dealt a three card hand consisting of three aces ?

Answer
Number of ways of selecting three aces (from the four in the pack) $=_4C_3$

$$\text{Thus probability} = \frac{_4C_3}{_{52}C_3}.$$

A special, but very common, case of the order of objects not being important (so that it is combinations we use rather than permutations) is when objects are identical and it is not possible to distinguish, never mind order, them!

Example 12
There are 10 balls in a bag, 6 red and 4 green. If 3 are picked out at random, what is the probability of 1 red and 2 greens ?

Answer
Total no of ways of picking 3 balls from 10 (regardless of colour) in which the order doesn't matter

$$= {}_{10}C_3 = \frac{10!}{7!3!} = 120$$

Outcomes: 3 reds   2 reds, 1 green   1 red, 2 greens   3 greens
The outcome we seek, 1 red and 2 greens, contains any 1 of the 6 reds and any 2 of the 4 greens. The number of ways it can come about is (since order doesn't matter)

$$= {}_6C_1 +_4 C_2 = \frac{6!}{5!1!} \frac{4!}{2!2!} = 36$$

$$\text{Thus probability} = \frac{36}{120} = \frac{3}{10}$$

As an exercise, we can check this by calculating the number of ways that the other outcomes can come about. (The total number, according to our answer, should be 120)

$$\text{3 reds} \qquad\qquad {}_6C_3 = \frac{6!}{3!3!} = 20$$

$$\text{2 reds, 1 green} \qquad\qquad {}_6C_2 \times_4 C_1 = \frac{6!}{4!2!} \frac{4!}{3!1!} = 60$$

$$\text{3 greens} \qquad\qquad {}_4C_3 = \frac{4!}{3!1!} = 4$$

$$\text{Total} = 36 + 20 + 60 + 4 = 120$$

Exercise 9

A mother takes her three young sons on a ride on the Severn Valley Railway. They get into an empty compartment with 3 forward-facing and 3 rearward-facing seats.

> (i) In how many different ways can they be seated ?
> (ii) In how many ways will two particular boys be seated directly opposite each other ?

Exercise 10

Three dice are thrown. Find the probability of obtaining

> (i) at least two sixes   (ii) no sixes
> (iii) different scores on all the dice.

Despite the small number of principles in the basic theory of probability, comparatively simple-looking problems can involve considerable subtlety in their solution. There is no substitute for experience and practice.

MEAN, VARIANCE AND STANDARD DEVIATION

Imagine a pack of cards with all the jokers and picture cards removed. We are only concerned with the numerical value of the cards. We have four each of all the numbers from one to ten so the pack contains 40 cards.

The average value, or arithmetic mean, of the numbers on the cards is given by adding them all up and dividing by the total number of cards. Let x_i be the number on the ith card and let N be the number of cards.

$$\text{Arithmetic mean: } \mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{220}{40} = 5.5$$

The arithmetic mean is a parameter of the set of numbers on the cards and it is normally referred to using the symbol μ. It is one measure of the *central tendency* or *middle value* of the numbers on the cards.

Another thing we might want to know is the *spread* of the values on the cards. One way to measure this is to calculate average of the squared differences between each value and the mean. This is called the variance.

$$\text{Variance: } \sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N} = \frac{330}{40} = 8.25$$

The standard deviation is another measure of spread. It is simply the square root of the variance. It is useful because it has the same dimensions as the original variable (i.e. the numbers on the cards).

$$\text{Standard Deviation: } \sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}} = \sqrt{\frac{330}{40}} = 2.8723$$

Example 13

It can be shown that for any set of data, at least $\left(1 - 1/k^2\right) \times 100\%$ of the data lies within k standard deviations of the mean.

For the population of numbers on the cards as described in the section above, check that this rule holds for $k = 1.2$

Answer

For $k = 1.2$ we have $\left(1 - 1/k^2\right) = 0.3056$ so we need to check that at least 30.56% of the data lies within 1.2 standard deviations of the mean.

In the previous section we calculated that the mean $\mu = 5.5$ and the standard deviation $\sigma = 2.8723$. The range within 1.2 standard deviations of the mean is therefore from 2.0533 to 8.9467, i.e. the numbers from 3 to 8 inclusive. This accounts for 24 cards, or 60% of the data. So more than the required 30.56% of the cards lie within 1.2 standard deviations of the mean.

Exercise 11

For each of the cases below, either find a set of five numbers that satisfies the required condition or explain why it cannot be achieved.

  (i) Four of the five numbers are below the mean.
 (ii) The mean is less than the standard deviation.
(iii) The mean is greater than the standard deviation.
(iv) The standard deviation is above the variance.

SAMPLE

Until now we have assumed that we can see all the cards at once. Now we are going to change the game. Imagine that someone else is holding the cards and allowing us to pick one at random, note its value and then replace it. Using this pick-and-replace process we can view a *sample* of the cards. This sample can be of any size as the cards are picked at random and replaced. Let us assume that the sample size is n.

The challenge is to estimate the mean, variance and standard deviation of the original numbers on the cards based only on what we see in the sample. Here are the formulae that enable us to do this.

$$\text{Estimate of Mean (based on sample): } m = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\text{Estimate of Variance (based on sample): } s^2 = \frac{\sum_{i=1}^{n}(x_i - m)^2}{n - 1}$$

$$\text{Estimate of Standard Deviation (based on sample): } s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n - 1}}$$

Notice that the formula for the mean is that same as that used in the previous section to calculate the mean of complete data set, however the formula to estimate the standard deviation is different. There is an $n - 1$ in the denominator rather than the N we had before. This is because we don't have the real mean, μ available so we have to use the *estimate* of the mean based on the sample, i.e. m. This is likely to be slightly better centred in the sample than the true mean (μ) of all the card values, hence the sum of the terms of the form $(x_i - m)^2$ will be slightly too small. The reduction in the denominator

from n to $n-1$ compensates for this. (Formal derivation of this result is beyond the scope of the first year course.)

There is a convenient formula for calculating the estimate of the standard deviation based on the sample.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n-1}} = \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}}$$

In the literature, s, the standard deviation of the underlying population estimated from a sample is called the *"sample standard deviation"*.

Example 14

Ten cards are selected individually, noted and replaced in the pack (that has no picture cards, as described above). This gives a sample size $n = 10$. The values of the cards are:

$$\begin{array}{ccccc} 10 & 3 & 4 & 3 & 5 \\ 4 & 1 & 5 & 8 & 5 \end{array}$$

Estimate the mean and standard deviation of the values in the complete pack based on this sample.

Answer

$$\text{Estimate of Mean: } m = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{48}{10} = 4.8$$

$$\text{Sample Standard Deviation: } s = \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}} = \sqrt{\frac{(290) - \frac{1}{10}(48)^2}{9}} = 2.5734$$

DISCRETE PROBABILITY DISTRIBUTION

Consider again the pack of cards with all the picture-cards removed. There are 40 cards each with an integer value in the range from one to ten. If we consider picking a card, noting its value and then replacing it in the pack we can compute the probability of picking each of the possible values. This

Probability
of value on
card

Value on the card
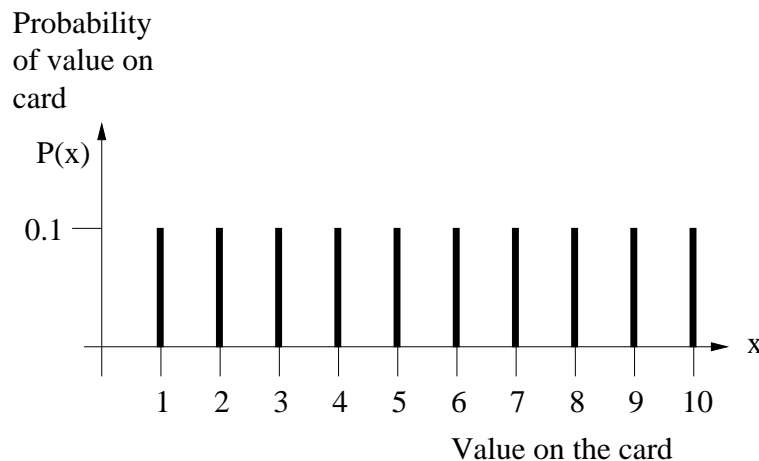
is a probability distribution. In this case it is a *discrete* distribution because the cards can only carry certain integer values. Notice that the sum of all the histogram bars is $10 \times 0.1 = 1$. There are ten possible outcomes and they each have a probability of 1/10. This is called a *uniform* distribution.

We can rearrange the formulae for the mean, variance and standard deviation, from the earlier section, to express these parameters in terms of the probability distribution, $P(x)$.

Note that the probability distribution is a property of the population of the numbers on the cards. Knowing the complete probability distribution enables us to calculate the mean μ and the standard deviation σ exactly. Let x_j represent each of the *different* values that are printed on the cards and M equal the number of these different values. In our example $M = 10$.

$$\text{Arithmetic mean: } \mu = \sum_{j=1}^{M} x_j P(x_j)$$

$$\text{Variance: } \sigma^2 = \sum_{j=1}^{M} (x_j - \mu)^2 P(x_j)$$

$$\text{Standard Deviation: } \sigma = \sqrt{\sum_{j=1}^{M} (x_j - \mu)^2 P(x_j)}$$

Example 15
Calculate the standard deviation of the numbers on the cards (as previously discussed), using the formulae based on the probability distribution.

Answer
First we need to calculate the mean. We can see from the histogram that $P(x_j) = 0.1$ for all the values on the cards (i.e. for all j). In this particular case, the values x_j are the same numerically as the index j, so we can substitute $x_j = j$. Hence

$$\mu = \sum_{j=1}^{M} x_j P(x_j) = \sum_{j=1}^{10} j \times 0.1 = 5.5$$

Now we use this value of μ in the formula for standard deviation.

$$\sigma = \sqrt{\sum_{j=1}^{M} (x_j - \mu)^2 P(x_j)} = \sqrt{\sum_{j=1}^{10} (j - 5.5)^2 \times 0.1} = 2.8723$$

Continuous Probability Distribution

We have seen how the diagram of the discrete probability distribution describes the probability of picking each value of card. There were 10 possible values so there were ten bars on the histogram.

Suppose however we had an random variable with 46 possible outcomes in the range 1–10. In this case we might get a discrete probability distribution such as that shown at the top of the next page on the left.

This histogram could perhaps describe the probability of being able to buy a particular electronic component at price of x pence, over a period of years when the cost of raw materials has varied. We can see that the manufacturers only choose prices which are multiples of 0.2 pence. The total area in all the histogram bars sums to 1. If we want to work out the probability that we might have got a price between 6 and 7 pence, we just have to add up the area of the histogram bars between 6 and 7.
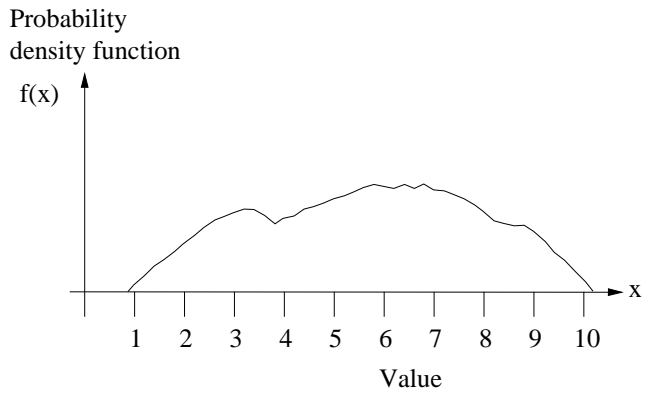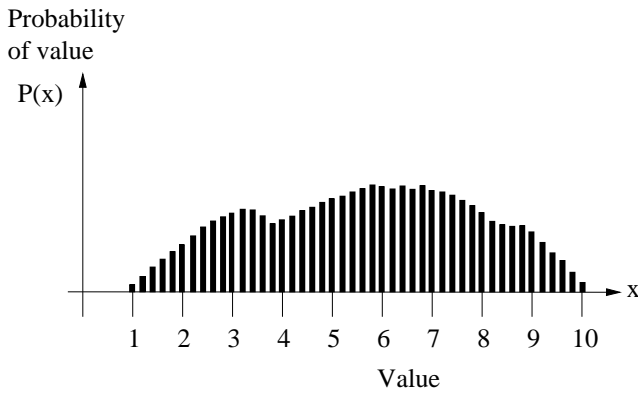
Suppose that the manufacturers of our electronic component did not choose to keep their prices as multiples of 0.2 pence and were happy to price them at any value. This would result in histogram bars that were so thin and close together that they would form a continuum. This leads to a continuous probability distribution. See the diagram above on the right.

Here the total area under the curve must be 1 and the probability of x taking a value in the range from (say) 6 to 7 is given by the integral (i.e. area) between 6 and 7. More generally:

$$\text{The probability of } (a < x < b) = \int_a^b f(x)\,dx$$

It is also possible to calculate the mean (μ) and standard deviation (σ) of a continuous distribution from its probability density function.

$$\mu = \int_{-\infty}^{+\infty} x f(x)\,dx \qquad \sigma = \sqrt{\int_{-\infty}^{+\infty} (x - \mu)^2 f(x)\,dx}$$

Example 16

Consider a machine that makes widgets which are supposed to be a particular length. Unfortunately, the machine often makes widgets that are slightly too long; it never makes widgets that are too short. The graph on the right shows the probability density function for the number of millimetres that a widget is too long.
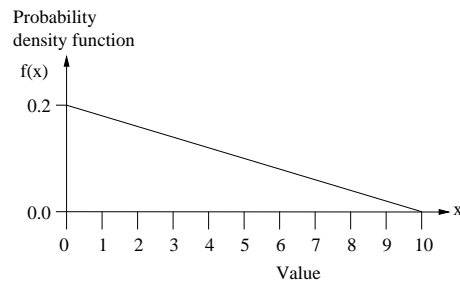
Calculate the mean and standard deviation of the excess length. What is the probability that a widget is produced with an excess length within one standard deviation from the mean.

Answer
For $0 \le x \le 10$ we can see that $f(x) = 0.2 - 0.02x$, hence:

$$\mu = \int_0^{10} x(0.2 - 0.02x)\,dx = 3.3333 \qquad \sigma = \sqrt{\int_0^{10} (x - 10/3)^2 (0.2 - 0.02x)\,dx} = 2.3570$$

So we wish to calculate the probability of an excess length in the range $3.33 - 2.36$ to $3.33 + 2.36$, which is given by:

$$P(x \text{ within one } \sigma \text{ of } \mu) = \int_{0.9763}^{5.6904} 0.2 - 0.02x\,dx = 0.6285$$

Continuous Normal Probability Distribution

Gaussian or Normal Probability Distribution with zero mean and unit standard deviation
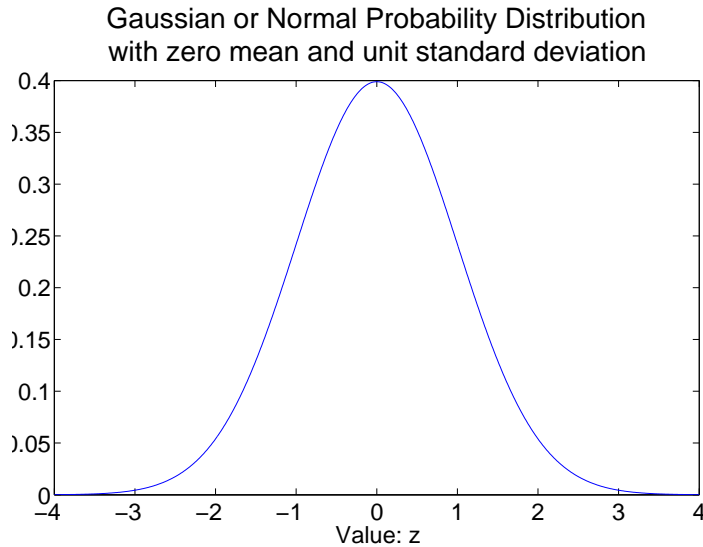
The Normal distribution is a symmetric distribution with two parameters, its mean μ and standard deviation σ. An example is illustrated here. It is a continuous distribution and the total area under the curve is 1. The formula for this curve is

$$P(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right)$$

although you do not have to use this directly very often in the first year.

It is difficult to evaluate integrals of the sort we did in Example 16 using the probability density function of the Normal distribution. Tables are therefore available and functions are built into Matlab and Octave to perform these calculations.

In general, if you add together a sufficient number (say 30), of independent random variables that are identically distributed, the sum will conform to the Normal distribution. This is called the "central limit theorem".

Important application: experimental data analysis

Consider a laboratory experiment that results in a single real output x, each time that we perform it. In theory, it should produce the same output each time but in practice x varies slightly because of noise in the measurement system.

If we repeat the experiment n times and average the result ($\overline{x} = m = \sum_{i=1}^{n} x_i/n$) then this average, $\overline{x}$, is itself a single sample from a random distribution. (If we ran the experiment a further n times then the average of these results would give us a second sample of $\overline{x}$ which would not, in general, be the same as the first.) We can say the following things about the distribution of $\overline{x}$:

- Provided $n > 30$ it is reasonable to assume that $\overline{x}$ is Normally distributed.

- The standard deviation of $\overline{x}$ will be a factor of $\sqrt{n}$ less than the standard deviation of the original experimental data. Hence:

$$\text{Estimate of standard deviation of experimental data} = \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}}$$

$$\text{Estimate of standard deviation of } \overline{x}: \quad s_{(\overline{x})} = \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}}$$

If it is fair to assume that the error in the original experimental data is unbiased, then the standard deviation of $\overline{x}$ gives us useful information about the error it is likely to contain. Let $s_{(\overline{x})}$ be our estimate of the standard deviation of $\overline{x}$.

50%	of the time, $\overline{x}$ will lie within	$0.67s_{(\overline{x})}$	of the true value.
68%	of the time, $\overline{x}$ will lie within	$s_{(\overline{x})}$	of the true value.
95%	of the time, $\overline{x}$ will lie within	$2s_{(\overline{x})}$	of the true value.
99.73%	of the time, $\overline{x}$ will lie within	$3s_{(\overline{x})}$	of the true value.

We can thus use the value of $s_{(\overline{x})}$ to calculate a range within which the true experimental result will lie with a particular probability, for example P(true result in range $\overline{x} \pm 2s_{(\overline{x})}$) = 0.95.

Example 17
An experiment is performed and a sequence of 30 readings are recorded.

2.8479	2.8613	3.2563	2.8381	2.7526	3.0429	3.4022	3.0051	3.0617	2.8124
3.3348	3.0250	3.1060	2.8096	3.1708	3.0778	2.7688	3.0079	2.9099	3.0218
2.9499	2.9620	2.7934	2.9353	3.1533	3.3489	2.7679	3.4755	3.3052	3.0337

The experimental error is believed to be unbiased. Calculate your best estimate of the true outcome of the experiment. Estimate the range within which there is a 95% chance that the true experimental result will lie.

Answer
Our estimate of the experimental outcome is the mean of these values $\overline{x} = 3.0279$.

Our estimate of the standard deviation of the experimental data is 0.2036, hence our estimate of the standard deviation of $\overline{x}$ is given by $0.2036/\sqrt{30} = 0.0372$.

We would expect the true experimental result to have a 95% chance of lying within two standard deviations from the mean, i.e. in the range 2.9536 to 3.1023.

Exercise 12
For the data in Example 17, estimate the range within which there is a 50% chance that the true experimental result will lie.

The remaining sections of this document are:-

APPENDIX A SOLUTIONS TO EXERCISES IN THE TEXT

**N.B.** $x!$, $_nC_r$ and $_nP_r$ appear as buttons on standard calculators.

1. (a) 26 of the 52 cards are red $\Rightarrow P(\text{red}) = \dfrac{26}{52} = \dfrac{1}{2}$

   (b) 4 of the 52 cards are kings $\Rightarrow P(\text{king}) = \dfrac{4}{52} = \dfrac{1}{13}$

   (c) 13 of the 52 cards are diamonds $\Rightarrow P(\text{diamond}) = \dfrac{13}{52} = \dfrac{1}{4}$

2. Outcomes are best listed as a table:-

1	1	1	2	1	3	1	4	1	5	1	6
2	1	2	2	2	3	2	4	2	5	2	6
3	1	3	2	3	3	3	4	3	5	3	6
4	1	4	2	4	3	4	4	4	5	4	6
5	1	5	2	5	3	5	4	5	5	5	6
6	1	6	2	6	3	6	4	6	5	6	6

There are 36 outcomes and 5 (5 1, 4 2, 3 3, 2 4, 1 5) result in a total of 6. $P(\text{total} = 6) = \dfrac{5}{36}$.

3. Let $E = set$ which pass the electrical and $M = pass\ mechanical$
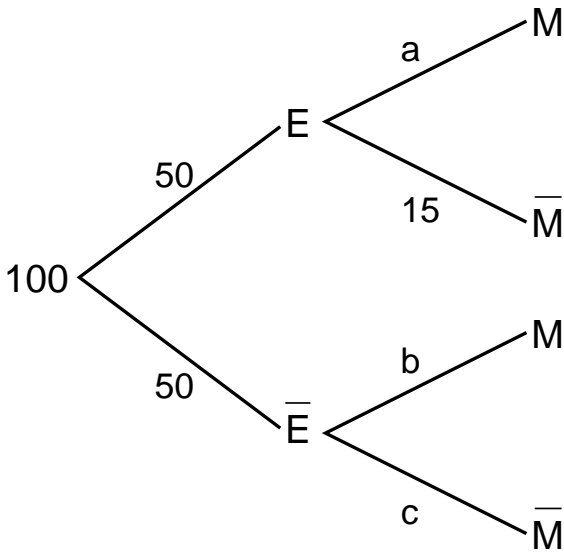
The data given is displayed on the branching diagram. Since there are 50 cars in group E, 35 must take branch a. Since 62 pass the mechanical test, then the number taking branch $b = 62 - 35 = 27$. Finally the number taking branch $c = 50 - 27 = 23$.

(i) This is a. Prob $= \dfrac{35}{100}$

(ii) Prob $= \dfrac{27}{100}(\text{branch } b) + \dfrac{15}{100}(M \cap \bar{E}) = \dfrac{42}{100}$.

(iii) Number failed = 50 (all of $\bar{E}$) + 15 ($M \cap \bar{E}$) = 65 . If a car has failed, it is one of these 65. 15 of these fail the mechanical test only. Therefore, required probability $= \dfrac{15}{65} = \dfrac{3}{13}$.

4. The events - obtain a six on die n - are *not exclusive* and instances of 6's occuring on more than one dice are counted (at least) twice by simply adding the probabilities.
Consider the probability of throwing a 6 with the first two dice. Since

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

P(at least one 6) = P(6 on 1st $\cup$ 6 on 2nd) = P(6 on 1st) + P(6 on 2nd) - P(6 on 1st $\cap$ 6 on 2nd)

$$= \frac{1}{6} + \frac{1}{6} - \frac{1}{6}\frac{1}{6} \text{ , etc.}$$
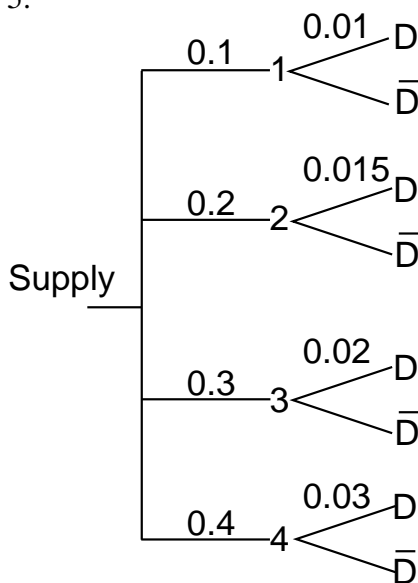
The correct argument is:

$$P(\text{ at least one 6 }) = 1 - P(\text{no sixes})$$

and throwing no six on each of the dice are *independent events* (probabilities multiply).

So P(no six on 1) $= \dfrac{5}{6} \Rightarrow$ P(no sixes) on 6 $= \left(\dfrac{5}{6}\right)^6 = 0.335$. P(at least one) $= 0.665$.

5.

The total proportion of the supply of bolts which is defective is
$0.1 \times 0.01 + 0.2 \times 0.015 + 0.3 \times 0.02$
$+0.3 \times 0.04 = 0.022$
The proportion of total supply which came from supplier 1 and is defective $0.1 \times 0.01 = 0.001$

P(Supplier 1 | defective) $= \dfrac{0.001}{0.022} = 0.045$

6. (i) $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow P(A \cap B) = \dfrac{1}{2} + \dfrac{1}{3} - \dfrac{2}{3} = \dfrac{1}{6}$

$P(A)P(B) = \dfrac{1}{2} \times \dfrac{1}{3} = \dfrac{1}{6}$ i.e. *independent* (since $P(A \cap B) = P(A)P(B)$)

(ii) $P(A \cap B) \neq 0 \Rightarrow$ *not* exclusive

7. (i) $\dfrac{8 \times 7 \times 6}{5 \times 4 \times 3} = \dfrac{8 \times 7 \times 6 \times 5 \times \ldots \times 1}{5 \times \ldots \times 1} \dfrac{2 \times 1}{5 \times \ldots \times 1} = \dfrac{8!2!}{5!5!}$

(ii) $\dfrac{15 \times 16}{4 \times 3 \times 2} = \dfrac{16!}{14!4!}$

(iii) $\dfrac{(n+1)n(n-1)}{4 \times 3 \times 2} = \dfrac{(n+1)!}{(n-2)!4!}$

$9 \times 7 \times 5 \times 3 \times 1 = \dfrac{9 \times 8 \times 7 \times 6 \times \ldots \times 1}{8 \times 6 \times 4 \times 2} = \dfrac{9!}{2^4 4!}$

22

8. The letters of CHEMISTRY form 9 distinct letters. If T is the last one, then the first 8 are rearrangements of CHEMISRY $\Rightarrow$ Number = 8! (8 choices for first, 7 for second, ...)

9. (i) 4 people to be arranged on 6 seats. The order matters, therefore, number $=_6 P_4 = 360$.
   (ii) 1 brother has 6 choices of seat. The position of the other one is then fixed.
   The remaining two members of the family select 2 from 4 (in which the order matters). The number of ways to do this is $_4P_2 = 12$.
   The total number of ways of seating the family in this manner $= 6 \times 12 = 72$.

10. (i) P(at least 2 sixes) = P(2 sixes) + P(3 sixes).

$$P(3 \text{ sixes}) = \left(\frac{1}{6}\right)^3 = 0.0046$$

P(2 sixes) = no of ways of choosing the dice which isn't the six
$\times$ probability it isn't a six $\times$ probability other 2 are sixes

$$= 3 \times \frac{5}{6} \times \left(\frac{1}{6}\right)^2 = 0.0694$$

So P(at least two sixes) = 0.074
Aliter:
Number of ways of throwing three sixes = 1
Number of ways of throwing 2 sixes =
        Number of ways of choosing the die which is not 6
        $\times$ Number of choices for a number which isn't six
        $= 3 \times 5$

Total Number of Throws $= 6^3$. Required Probability $= \dfrac{1 + 3 \times 5}{6^3}$.

(ii) P(no sixes) $= \left(\dfrac{5}{6}\right)^3 = 0.58$

(iii) 6 choices for first, 5 for second, 4 for third = 120 ($=_6 P_3$),

Probability $= \dfrac{120}{6^3} = 0.56$.

11(i) In this case we need four small numbers and a single large number to give a high overall mean. For example: 1, 1, 1, 1, 46 which have a mean of 10.
  (ii) The mean can take on any value, large or small, because the numbers can be negative. The population: $-3, -1, 0, 1, 3$ has a mean of zero and a standard deviation of 2.
 (iii) We can produce a population with a mean greater than the standard deviation by using the answer to the previous part and adding 3 to each number. This will result in a mean of $0 + 3 = 3$ and the same standard deviation as before. The answer is therefore: 0, 2, 3, 4, 6 which has $\mu = 3$ and $\sigma = 2$.
 (iv) The standard deviation is the square root of the variance and both are always non-negative. To get $\sigma > \sigma^2$ we require $\sigma < 1$. A suitable population is therefore: $-1, -1, 0, 1, 1$ which has variance $\sigma^2 = 0.8$ and standard deviation $\sigma = 0.8944$.

12. For the data in Example 17 the standard deviation of the distribution of $\bar{x}$ is given by $s_{(\bar{x})} = 0.0372$. For 50% of the time, the true value will be within $0.67 s_{(\bar{x})}$ of the mean $\bar{x} = 3.0279$. Hence we want the range $\bar{x} - 0.67 s_{(\bar{x})}$ to $\bar{x} + 0.67 s_{(\bar{x})}$ which works out to be 3.0030 to 3.0528.

APPENDIX B   FURTHER WORKED EXAMPLES

The questions in this section can be used as worked examples, in which case you should read through them carefully. They can also be used as a set of further exercises, in which case attempt each question, or at least think about how you would go about solving it, before working through the solution.

N.B. $x!$, $_nC_r$ and $_nP_r$ appear as buttons on standard calculators.

---

1.   A well shuffled pack of 52 playing cards is dealt out to four players, each receiving thirteen cards. Show that the probability that a particular player receives the four aces is 0.0026. How many deals are necessary in order that the probability of a particular player receiving all four aces at least once exceeds 0.5 ?

Answer

52 cards   =   4 aces   +   48 others
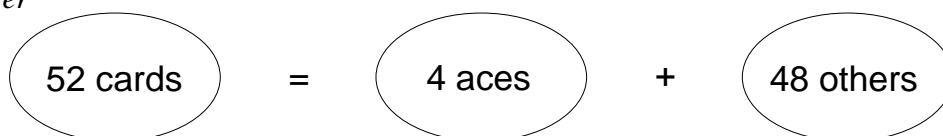
Player with 4 aces has 9 cards from amongst the other 48.

$$\Rightarrow \text{P(4 aces)} = \frac{_{48}C_9}{_{52}C_{13}} = 0.00264$$

The probability of not getting the 4 aces in one deal is $1 - 0.00264 = 0.99736$
Probability of not getting 4 aces in N deals (considered independent) $= 0.99736^N$
Probability of getting 4 aces at least once in N deals $= 1 - 0.99736^N$
To ensure that this is greater than 0.5, we need

$$1 - 0.99736^N > 0.5$$
$$\Rightarrow N\log(0.99736) < \log(0.5) \Rightarrow N > 262.2 \text{ i.e. } N = 263$$

---

2.   Find how many teams of two pairs (1st and 2nd), each pair consisting of a man and a woman, can be chosen to represent a club at a badminton match, if these pairs are chosen from five men and three women.

Answer
The team is $M_1 \ W_1 : M_2 \ W_2$.
To select the men, we need to select two from the five available and the *order matters*. This can be done in $_5P_2$ ways. The women can similarly be selected in $_3P_2$. If the selection of men and women is independent, then the total number of ways of selecting the team is

$$_5P_2 \times _3P_2 = 20 \times 6 = 120$$

---

3.   My parents are celebrating their golden wedding anniversary and we are going to have a big family dinner for them at home. The dining-table will seat only 12, but there are 17 members of the family (including us) who live near enough to be invited. Ignoring any restrictions on choice, in how many ways can I invite 12 members of the family?

Of course, I must invite my parents, and my brother and I must be there; in how many ways can I now invite 12 ?

Two of the 17 are the dreaded Aunty Joyce and her husband Reg. If I invite her, I must invite him also. Taking this into account, in how many ways can I now invite everyone ?

Having finally chosen the 12 who will be present at the dinner, they now have to be seated at the large round dinner table. I have to be seated nearest to the kitchen. In how many ways can we all sit down to dinner?

After much debate we have decided to invite Reg and Joyce. They are such a nuisance that it would be best to sit them down together (but Reg must not be next to me!). In how many ways can the party now be seated ?

Answer

Order doesn't matter. Therefore 12 can be invited in $_{17}C_{12} = \underline{6188}$.

Four individuals are determined and 8 from the other 13 can come. Number of ways $=_{13}C_8 = \underline{1287}$.

If invite AJ & UR, remaining places filled in $_{11}C_6 = 462$.
If don't invite AJ & UR, remaining places filled in $_{11}C_8 = 165$. Thus total $= 462 + 165 = \underline{627}$

Table is round. I sit nearest to the kitchen. The other 11 can be seated in $\underline{11!}$ ways (order matters).
Some people would say $\frac{11!}{2}$ (if mirror image seating plans are counted as the same).

My seat determined, 9 choices for UR, 2 choices for AJ, then 9! ways for the rest. i.e. $\underline{18 \times 9!}$
(or $9 \times 9!$ if mirror images not counted as different).

---

4.   In a game of bridge the pack of 52 cards is shared equally between four players. What is the probability that a particular player has no picture cards (Jacks, Queens, Kings)?

Answer

The total number of hands that the player can hold = the number of ways of choosing 13 cards from 52 (in which the order does not matter) $=_{52}C_{13}$.

Now

| 52 cards | = | 40 non-picture cards | + | 12 Jacks Queens & Kings |

If the player has no picture cards then the cards dealt must all be chosen from the group of 40. This can be done in $_{40}C_{13}$ ways. The probability that a particular player has no picture cards is thus

$$\frac{\text{Number of deals with no pictures}}{\text{Number of deals}} = \frac{_{40}C_{13}}{_{52}C_{13}} = \frac{1.20 \times 10^{10}}{6.35 \times 10^{11}} = 0.019$$

---

5.   I have a box of chocolates with 10 different chocolates in it. Of these, there are 6 which I particularly like. However, I intend to offer my three friends one chocolate each before I eat the rest. How many different selections of chocolates can I be left with after my friends have chosen ?
Show that 36 of these selections leave me with exactly 5 chocolates which I particularly like.
How many selections leave me with (i) all 6; (ii) exactly 4; and (iii) exactly 3 of the chocolates which I particularly like?
Assuming that my friends choose at random, what is the most likely outcome and what is the probability of that outcome?

Answer
After friends have chosen, there are 7 out of the original 10 left. Order doesn't matter. Therefore the number of selections $=_{10}C_7 = 120$.

| 7 | = | 3 I like | + | 4 | or | 4 I like | + | 3 | or | 5 I like | + | 2 | or | 6 I like | + | 1 |

Number of ways of choosing these are

$$_6C_3\ _4C_4 \quad \text{or} \quad _6C_4\ _4C_3 \quad \text{or} \quad _6C_5\ _4C_2 \quad \text{or} \quad _6C_6\ _4C_1$$
$$20 \qquad\qquad\quad 60 \qquad\qquad\quad 36 \qquad\qquad\quad 4$$

(i) all six left - 4 ways   (ii) four left - 60 ways   (iii) three left - 20

Most likely outcome $= 4$ left and $P(4) = \dfrac{60}{120} = \dfrac{1}{2}$.

---

6   A multiple-choice exam has 10 questions. Each question has 3 possible 'answers' of which only 1 is correct. A candidate attempts the exam using pure guesswork. Find the probabilities that the candidate obtains
(i) exactly 2 correct answers,   (ii) not more than 2 correct answers.
What is the most likely number of correct answers and the probability that the candidate just achieves this number ?

Answer
(i) The 2 correct questions can be chosen from the 10 in $_{10}C_2$ ways.

The probability that these two questions will be answered correctly $= \left(\dfrac{1}{3}\right)^2$

The probability that the other eight will be incorrect $= \left(\dfrac{2}{3}\right)^8$.

The probability of getting these 2 right and the other 8 wrong, therefore, $= \left(\dfrac{1}{3}\right)^2\left(\dfrac{2}{3}\right)^8$ assuming the guesses for the answers are made independently (hence multiply the probabilities). This probability is appropriate for each possible selection of two questions (and each selection is an exclusive event, therefore add the probabilities). The required probability is, therefore,

$$_{10}C_2\left(\dfrac{1}{3}\right)^2\left(\dfrac{2}{3}\right)^8 = 45 \times \dfrac{1}{9} \times \dfrac{256}{6561} = 0.195$$

(ii) If not more than 2 are answered correctly, number correct can be 0, 1 or 2. Required probability is, by exactly similar reasoning,

$$_{10}C_0\left(\dfrac{2}{3}\right)^{10} + _{10}C_1\left(\dfrac{1}{3}\right)^1\left(\dfrac{2}{3}\right)^9 + _{10}C_2\left(\dfrac{1}{3}\right)^2\left(\dfrac{2}{3}\right)^8 = 0.017 + 0.087 + 0.195 = 0.299$$

(iii) Probability of getting r correct $= P(r) =_{10} C_r \left(\frac{1}{3}\right)^r \left(\frac{2}{3}\right)^{10-r}$

This increases with r all the time that $P(r+1) > P(r)$.

$$\text{Now } \frac{P(r+1)}{P(r)} = \frac{_{10}C_{r+1}\left(\frac{1}{3}\right)^{r+1}\left(\frac{2}{3}\right)^{10-r-1}}{_{10}C_r\left(\frac{1}{3}\right)^r\left(\frac{2}{3}\right)^{10-r}} = \frac{10!}{(10-r-1)!(r+1)!} \frac{r!(10-r)!}{10!} \frac{1}{3}\left(\frac{2}{3}\right)^{-1}$$

$$= \frac{10-r}{r+1}\frac{1}{2}.$$

Thus $P(r+1) > P(r)$ until $\dfrac{10-r}{r+1}\dfrac{1}{2} < 1$ i.e. $10-r < 2r+2 \Rightarrow r > 2.67$.

Thus $P(2+1) > P(2)$, but $P(3+1) < P(3)$. $r = 3$ is thus most likely.

Check $P(3) =_{10} C_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^7 = 0.26$ $P(4) =_{10} C_4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^6 = 0.23$

Engineering FIRST YEAR

Part 1A Paper 4: Mathematics

Examples Paper 9 - Probability

(Elementary exercises are marked †, problems of Tripos standard ∗)

N.B. $x!$, $_nC_r$ and $_nP_r$ appear as buttons on standard calculators.

1† Find the number of integers between 1000 and 4000 which can be formed using the digits 1,2,3,4

        (i) if each digit may be used only once,

        (ii) if each digit may be used more than once.

2† How many odd numbers can be formed from the digits 2,3,4,5,6 and 7 using each digit once and once only ?

3† The probability that a January night will be icy in Cambridge is 0.25. On an icy night the probability of an accident on Castle Hill is 0.04, and 0.01 on non-icy nights. What is the probability that

        (i) 13th January will be icy and there will be an accident on Castle Hill

        (ii) there will be an accident on 13th January.

4 Show that the probability that a point, selected at random along a line of unit length, is closer to the centre of the line than to either of the ends is $\frac{1}{2}$.

Show that the probability that a point, selected at random inside a circle, is closer to the centre of the circle than to the circumference is $\frac{1}{4}$.

Points are selected at random inside the circle until a point is closer to the centre than to the circumference. What is the probability that

        (a) exactly three points are selected

        (b) no more than three points are selected ?

How many points need to be selected so that there is a probability of at least 0.85 that at least one point is closer to the centre than to the circumference ?

5 In Camelot it never rains on Friday, Saturday, Sunday or Monday. The probability that it rains on a given Tuesday is $\frac{1}{5}$. For each of the remaining two days, Wednesday and Thursday, the conditional probability that it rains, given that it rained the previous day, is α, and the conditional probability that it rains, given that it did not rain the previous day, is β .

(a) Show that the (unconditional) probability of rain on a given Wednesday is $\frac{1}{5}[\alpha + 4\beta]$, and find the probability of rain on a given Thursday.

(b) If X is the event that, in a randomly chosen week, it rains on Thursday, Y is the event that it rains on Tuesday, and $\bar{Y}$ is the event that it does not rain on Tuesday, show that

$$P(X|Y) - P(X|\bar{Y}) = (\alpha - \beta)^2.$$

(c) Explain the implications of the case $\alpha = \beta$.

6 For the game of bridge each of 4 players gets 13 cards from a pack of 52. Taking account of the different ways in which the hands can be distributed amongst the 4 players, how many different deals are there ?

7 Show that the probability of drawing 13 cards at random from a pack of 52 and having no spades in that hand is 0.0128. Explain why the probability of having at least one suit missing is approximately 4 times this.

8† Two students have five coins and three coins respectively. They while away their time waiting for a lecture by tossing them all once. What is the probability that they will both toss the same number of tails ?

9† In a well-known game you have to decide which your opponent is going to choose: 'Paper', 'Stone' or 'Scissors'. If you guess entirely at random, what is the probability that you are right exactly 5 times out of 15 ?

10 A coin is tossed ten times.

> (i) What is the probability of it coming down heads 5 times and tails 5 times ?
>
> (ii) Which is the more likely: exactly 7 heads or more than 7 heads ?

Verify your answers by tossing a coin ten times and counting the number of heads. Repeat the experiment 100000 times and record the frequencies of the various numbers of heads. If you do not have time for 100000 repetitions, you could alternatively write a Matlab/Octave program to simulate the experiments.[1]

11 A certain strain of rabbit produces offspring in equal ratio male to female. One third of the male offspring and one half of the female offspring are black. In a litter of 5 offspring, what is the probability that

> (i) exactly 2 offspring will be male;
>
> (ii) less than 4 offspring will be female;

In a litter of 4 offspring, what is the probability that there will be exactly 2 male offspring which are not black ?

12∗ A bag P contains 3 red balls. A second bag Q contains 2 red balls and 3 black balls.

> (a) A bag is chosen at random and one ball is withdrawn. Find the probability that this ball is red.
>
> This ball remains outside the bag.
>
> (b) A bag is again chosen at random (it is not known whether this is the same bag as before or not) and one ball is withdrawn. Find the joint probability that both this ball and the one previously withdrawn are red.

---

[1]Hint: code for doing this can be downloaded from `http://sigproc.eng.cam.ac.uk/Main/1AMathematics_2013`. The code comes in a file called q10.m. Save this in a folder somewhere, start Matlab/Octave from the same folder (or use the "cd" command to navigate to that folder), then type "q10" to run the code. Use a text editor to change the number of experiments near the top of the file q10.m, then run the code again.

(c) If they are both red, what is the probability that bag P was used on both occasions ?

13 Consider a special pack of ten cards contains the following values: one ace, two twos, three threes and four fours.

(i) What is the mean value of this population of ten values?

(ii) What is the standard deviation of this population of ten values?

(iii) Twelve times I pick a card randomly from this pack, note its value and replace it. I get the following values: $3, 4, 4, 2, 3, 1, 3, 3, 4, 4, 3, 4$. Use this sample to estimate the mean and standard deviation of the population of values on the cards.

14 An experiment is performed 100 times and separate values x_i are recorded for i equals 1 to 100. The researcher reports that $\sum_{i=1}^{100} x_i = 25,545.1678$ and $\sum_{i=1}^{100} x_i^2 = 6,530,839.6114$ Estimate a range in which you are 99.73% sure that the true experimental result lies.

Answers

1 (i) 18   (ii) 192

2 360

3 (i) 0.01   (ii) 0.0175

4 (a) $\dfrac{9}{64}$ (b)$\dfrac{37}{64}$ 7 points

5 (a) $\beta + \dfrac{(\alpha - \beta)(\alpha + 4\beta)}{5}$

6 5.4×10^{28}

8 $\dfrac{7}{32}$

9 0.214

10 (i) 0.246   (ii) 7 heads

11 (i) 0.31   (ii) 0.81   0.30

12 (a) $\dfrac{7}{10}$ (b) $\dfrac{19}{40}$ (c) $\dfrac{10}{19}$

13 (i) 3, (ii) 1, (iii) 3.1667, 0.9374

14 253.26 to 257.64.