Chapter 2. Counting techniques

Bibliografie: Riley et al. (2006)

In the previous chapter we defined the probability of an event *A* in a sample space *S* as

$$\Pr(A) = \frac{n_A}{n_S} \tag{2.1}$$

where n_A is the number of outcomes belonging to event *A* and n_S is the total number of possible outcomes. It is therefore necessary to be able to count the number of possible outcomes in various common situations.

2.1 Permutations

Let us first consider a set of *n* objects that are all different. We may ask in how many ways these *n* objects may be arranged, i.e. how many permutations of these objects exist. This is straightforward to deduce: the object in the first position may be chosen in *n* different ways, that in the second position in $n-1$ ways, and so on until the final object is positioned. The number of possible arrangements is therefore

$$n(n-1)(n-2)\ldots 2\times 1 = n! \tag{2.2}$$

Let us suppose we choose only $k(<n)$ objects from *n*. The number of possible permutations of these *k* objects selected from *n* is given by

$$\underbrace{n(n-1)(n-2)\ldots(n-k+1)}_{k\ \text{factors}} = \frac{n!}{(n-k)!} = A_n^k \tag{2.3}$$

In calculating the number of permutations of the various objects we have so far assumed that the objects are sampled *without replacement* – i.e. once an object has been drawn from the set it is put aside. However, we may instead replace each object before the next is chosen. The number of permutations of *k* objects from *n* *with replacement* may be calculated very easily since the first object can be chosen in *n* different ways, as can the second, the third, etc. Therefore the number of permutations is simply n^k. This may also be viewed as the number of permutations of k objects from n where repetitions are allowed, i.e. each object may be used as often as one likes.

Exercise 1. Find the probability that in a group of k people at least two have the same birthday (ignoring 29 February).

It is simplest to begin by calculating the probability that no two people share a birthday, as follows. Firstly, we imagine each of the k people in turn pointing to their birthday on a year planner. Thus, we are sampling the 365 days of the year 'with replacement' and so the total number of possible outcomes is $(365)^k$. Now (for the moment) we assume that no two people share a birthday and imagine the process being repeated, except that as each person points out their birthday it is crossed off the planner. In this case, we are sampling the days of the year 'without replacement', and so the possible number of outcomes for which all the birthdays are different is

$$A_{365}^{k} = \frac{365!}{(365-k)!}$$

Hence the probability that all the birthdays are different is

$$p = \frac{365!}{(365-k)!365^k}$$

Using the complement rule, the probability q that two or more people have the same birthday is simply

$$q = 1 - p = 1 - \frac{365!}{(365-k)!365^k}$$

This expression may be conveniently evaluated using Stirling's approximation for $n!$ when n is large, namely

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$$365! \approx \sqrt{2\pi 365}\left(\frac{365}{e}\right)^{365}$$

$$(365-k)! \approx \sqrt{2\pi(365-k)}\left(\frac{365-k}{e}\right)^{365-k}$$

$$p=\frac{365!}{(365-k)!365^k}=\frac{\sqrt{2\pi 365}}{\sqrt{2\pi(365-k)}}\left(\frac{365}{e}\right)^{365}\left(\frac{e}{365-k}\right)^{365-k}\frac{1}{365^k}$$

$$=\sqrt{\frac{365}{365-k}}\left(\frac{365}{365-k}\right)^{365-k}e^{-k}=e^{-k}\left(\frac{365}{365-k}\right)^{365-k+0.5}$$

$$q\approx 1-e^{-k}\left(\frac{365}{365-k}\right)^{365-k+0.5}$$

If $k=23$ the probability is a little greater than $1/2$ that at least two people have the same birthday, and if $k=50$ the probability rises to 0.970.

So far we have assumed that all *n* objects are different (or *distinguishable*). Let us now consider *n* objects of which n_1 are identical and of type 1, n_2 are identical and of type 2, . . . , n_m are identical and of type *m* ($n=n_1+n_2+\ldots+n_m$). From (2.2) the number of permutations of these *n* objects is again $n!$. However, the number of *distinguishable* permutations is only

$$\frac{n!}{n_1!n_2!\cdots n_m!} \tag{2.4}$$

since the *i*-th group of identical objects can be rearranged in $n_i!$ ways without changing the distinguishable permutation.

Exercise 2. A set of snooker balls consists of a white, a yellow, a green, a brown, a blue, a pink, a black and 15 reds. How many distinguishable permutations of the balls are there?

In total there are 22 balls, the 15 reds being indistinguishable. Thus from (2.4) the number of distinguishable permutations is

$$\frac{22!}{(1!)(1!)(1!)(1!)(1!)(1!)(1!)(15!)}=\frac{22!}{15!}=859\ 541\ 760$$

2.2 Combinations

We now consider the number of combinations of various objects when their order is immaterial. Assuming all the objects to be distinguishable, the number of permutations of *k* objects chosen from *n* is $A_n^k = n!/(n-k)!$.

Now, since we are no longer concerned with the order of the chosen objects, which can be internally arranged in $k!$ different ways, the number of combinations of *k* objects from *n* is

$$\frac{n!}{(n-k)!k!} = C_n^k \quad \text{for} \quad 0 \le k \le n \tag{2.5}$$

C_n^k is called the *binomial coefficient* since it also appears in the binomial expansion

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k} \tag{2.6}$$

Exercise 3. A hand of 13 playing cards is dealt from a well-shuffled pack of 52. What is the probability that the hand contains two aces?

Since the order of the cards in the hand is immaterial, the total number of distinct hands is simply equal to the number of combinations of 13 objects drawn from 52, i.e. C_{52}^{13}. However, the number of hands containing two aces is equal to the number of ways, C_4^2, in which the two aces can be drawn from the four available, <u>multiplied</u> by the number of ways, C_{48}^{11}, in which the remaining 11 cards in the hand can be drawn from the 48 cards that are not aces. Thus the required probability is given by

$$\frac{C_4^2 C_{48}^{11}}{C_{52}^{13}} = \frac{4!}{2!2!} \times \frac{48!}{11!37!} \times \frac{13!39!}{52!} = 0.213$$

BASIC PRINCIPLE OF COUNTING: Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of *m* possible outcomes and if, for each outcome of experiment 1, there are *n* possible outcomes of experiment 2, then together there are mn possible outcomes of the two experiments.

Another useful result is the number of ways in which *n* distinguishable objects can be divided into *m* piles, with n_i objects in the *i*-th pile, $i = 1, 2, \ldots, m$ (the ordering of objects within each pile being unimportant). We may choose the n_1 objects in the first pile from the original *n* objects in $C_n^{n_1}$ ways. The n_2 objects in the second pile can then be chosen from the $n - n_1$ remaining objects in $C_{n-n_1}^{n_2}$ ways, etc. We may continue in this fashion until we reach the $(m-1)$-th pile, which may be formed in $C_{n-n_1-\ldots-n_{m-2}}^{n_{m-1}}$ ways. The remaining objects then form the *m*-th pile and so can only be 'chosen' in one way. Thus the total number of ways of dividing the original *n* objects into *m* piles is given by the product

$$N = C_n^{n_1} C_{n-n_1}^{n_2} \ldots C_{n-n_1-\ldots-n_{m-2}}^{n_{m-1}} \cdot 1 =$$

$$= \frac{n!}{n_1!(n-n_1)!} \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!} \ldots \frac{(n-n_1-\ldots n_{m-2})!}{n_{m-1}!(n-n_1-\ldots n_{m-2}-n_{m-1})!}$$

$$= \frac{n!}{n_1!n_2!n_3!\ldots(n_{m-1})!(n_m)!} \qquad (2.7)$$

These numbers are called *multinomial coefficients* since (2.7) is the coefficient of $x_1^{n_1} x_2^{n_2} \ldots x_m^{n_m}$ in the multinomial expansion of $(x_1 + x_2 + \ldots + x_m)^n$. Furthermore, we note that the multinomial coefficient (2.7) is identical to the expression (2.4) for the number of distinguishable permutations of *n* objects, n_i of which are identical and of type *i* (for $i = 1, 2, \ldots, m$ and $n = n_1 + n_2 + \ldots + n_m$).

Exercise 4. In the card game of bridge, each of four players is dealt 13 cards from a full pack of 52. What is the probability that each player is dealt an ace?

From (2.7), the total number of distinct bridge dealings is $\frac{52!}{13!13!13!13!}$ ($n = 52$ cards, 4 piles with 13 cards). The number of ways in which the four aces can be distributed with one in each hand is $\frac{4!}{1!1!1!1!} = 4!$; the remaining 48 cards can be dealt out in $\frac{48!}{12!12!12!12!}$ ways. Thus the probability that each player receives an ace is

$$4! \frac{48!}{(12!)^4} / \frac{52!}{(13!)^4} = 0.105$$

As in the case of permutations we might ask how many combinations of k objects can be chosen from n with replacement (repetition). To calculate this, we may imagine the n (distinguishable) objects set out on a table. Each combination of k objects can then be made by pointing to k of the n objects in turn (with repetitions allowed). These k equivalent selections distributed amongst n different but re-choosable objects are strictly analogous to the placing of k indistinguishable 'balls' in n different boxes with no restriction on the number of balls in each box. A particular selection in the case $k=7$, $n=5$ may be symbolised as

$$xxx|\;|x|xx|x$$

This denotes three balls in the first box, none in the second, one in the third, two in the fourth and one in the fifth. We therefore need only consider the number of (distinguishable) ways in which k crosses and $n-1$ vertical lines can be arranged, i.e. the number of permutations of $k+n-1$ objects of which k are identical crosses and $n-1$ are identical lines. This is given by (2.4) as

$$\frac{(k+n-1)!}{k!(n-1)!}=C_{n+k-1}^{k} \tag{2.8}$$

Exercise 5. A system contains a number N of (non-interacting) particles, each of which can be in any of the quantum states of the system. The structure of the set of quantum states is such that there exist R energy levels with corresponding energies E_i and degeneracies g_i (i.e. the ***i***-th energy level contains g_i quantum states). Find the numbers of distinct ways in which the particles can be distributed among the quantum states of the system such that the ***i***-th energy level contains n_i particles, for $i=1,2,\ldots,R$, in the cases where the particles are:

(i) distinguishable with no restriction on the number in each state;

(ii) indistinguishable with no restriction on the number in each state;

(iii) indistinguishable with a maximum of one particle in each state;

(iv) distinguishable with a maximum of one particle in each state.

It is easiest to solve this problem in two stages. First consider distributing the *N* particles among the *R* energy levels, *without* regard for the individual degenerate quantum states that comprise each level. If the particles are *distinguishable* then the number of distinct arrangements with n_i particles in the *i*-th level, $i = 1, 2, \ldots, R$, is given by (2.7) as

$$\frac{N!}{n_1! n_2! \cdots n_R!}$$

If the particles are *indistinguishable* then clearly there exists only one distinct arrangement having n_i particles in the *i*-th level, $i = 1, 2, \ldots, R$. If we suppose that there exist w_i ways in which the n_i particles in the *i*-th energy level can be distributed among the g_i degenerate states, then it follows that the number of distinct ways in which the *N* particles can be distributed among all *R* energy levels of the system, with n_i particles in the *i*-th level, is given by

$$W\{n_i\} = \begin{cases} \dfrac{N!}{n_1! n_2! \cdots n_R!} \displaystyle\prod_{i=1}^{R} w_i & \text{for distinguishable particles} \\ \displaystyle\prod_{i=1}^{R} w_i & \text{for indistinguishable particles} \end{cases} \tag{2.9}$$

It therefore remains only to find the appropriate expression for w_i in each of the cases (i)–(iv) above.

Case (i). If there is no restriction on the number of particles in each quantum state, then in the *i*-th energy level each particle can reside in any of the g_i degenerate quantum states. Thus, if the particles are distinguishable then the number of distinct arrangements is simply $w_i = g_i^{n_i}$. Thus, from (2.9),

$$W\{n_i\} = \frac{N!}{n_1! n_2! \cdots n_R!} \prod_{i=1}^{R} g_i^{n_i} = N! \prod_{i=1}^{R} \frac{g_i^{n_i}}{n_i!}$$

Such a system of particles (for example atoms or molecules in a classical gas) is said to obey *Maxwell–Boltzmann statistics*.

Case (ii). If the particles are *indistinguishable* and there is *no restriction* on the number in each state then, from (2.8), the number of distinct arrangements of the n_i particles among the g_i states in the *i*-th energy level is

$$w_i = \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}$$

Substituting this expression in (2.9), we obtain

$$W\{n_i\} = \prod_{i=1}^{R} \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}$$

Such a system of particles (for example a gas of photons) is said to obey *Bose–Einstein statistics*.

Case (iii). If a maximum of one particle can reside in each of the g_i degenerate quantum states in the ***i***-th energy level then the number of particles in each state is either 0 or 1. Since the particles are *indistinguishable*, w_i is equal to the number of distinct arrangements in which n_i states are occupied and $g_i - n_i$ states are unoccupied; this is given by

$$w_i = C_{g_i}^{n_i} = \frac{g_i!}{n_i!(g_i - n_i)!}$$

Thus, from (2.9), we have

$$W\{n_i\} = \prod_{i=1}^{R} \frac{g_i!}{n_i!(g_i - n_i)!}$$

Such a system is said to obey *Fermi–Dirac statistics*, and an example is provided by an electron gas.

Case (iv). Again, the number of particles in each state is either 0 or 1. If the particles are *distinguishable*, each arrangement identified in case (iii) can be reordered in $n_i!$ different ways, so that

$$w_i = A_{g_i}^{n_i} = \frac{g_i!}{(g_i - n_i)!}$$

Substituting this expression into (2.9) gives

$$W\{n_i\} = \frac{N!}{n_1!n_2!\cdots n_R!} \prod_{i=1}^{R} \frac{g_i!}{(g_i - n_i)!} = N! \prod_{i=1}^{R} \frac{g_i!}{n_i!(g_i - n_i)!}$$

Such a system of particles has the names of no famous scientists attached to it, since it appears that it never occurs in nature.

Chapter 3 Random variables and distributions

Bibliografie: Riley et al. (2006)

Suppose an experiment has an outcome sample space *S*. A real variable *X* that is defined for all possible outcomes in S (so that a real number is assigned to each possible outcome) is called a *random variable* (RV).

$$X : S \to \mathbb{R}$$

The outcome of the experiment may already be a real number and hence a random variable, e.g. the number of heads obtained in 10 throws of a coin, or the sum of the values if two dice are thrown. However, more arbitrary assignments are possible, e.g. the assignment of a 'quality' rating to each successive item produced by a manufacturing process.

Assuming that a probability can be assigned to all possible outcomes in a sample space *S*, it is possible to assign a *probability distribution* to any random variable. Random variables may be divided into two classes, discrete and continuous, and we now examine each of these in turn.

3.1 Discrete random variables

A random variable *X* that takes only discrete values $x_1, x_2, \dots, x_n$ with probabilities $p_1, p_2, \dots, p_n$ is called a *discrete random variable*. The number of values *n* for which *X* has a non-zero probability is finite or at most countably infinite. An example of a discrete random variable is the number of heads obtained in 10 throws of a coin. If *X* is a discrete random variable, we can define a *probability function* (PF) $f(x)$ that assigns probabilities to all the distinct values that *X* can take, such that

$$f(x) = \Pr(X = x) = \begin{cases} p_i \text{ if } x = x_i \\ 0 \text{ otherwise} \end{cases} \tag{3.1}$$

A typical PF (see figure 3.1) thus consists of spikes, at valid values of *X*, whose height at *x* corresponds to the probability that $X = x$. Since the probabilities must sum to unity, we require

$$\sum_{i=1}^{n} f(x_i) = 1 \tag{3.2}$$

We may also define the *cumulative probability function* (CPF) of *X*, $F(x)$, whose value gives the probability that $X \le x$, so that

$$F(x) = \Pr(X \le x) = \sum_{x_i \le x} f(x_i) \tag{3.3}$$

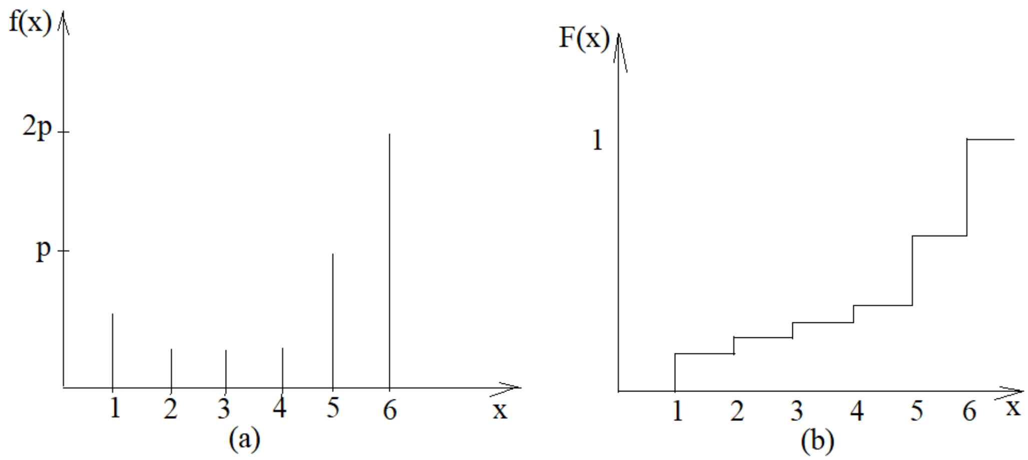

Figura 3.1 (a) A typical probability function for a discrete distribution, that for the biased die discussed earlier (exercise 4, chapter 1). (b) The cumulative probability function for the same discrete distribution.

Hence $F(x)$ is a step function that has upward jumps of p_i at $x = x_i$, $i = 1, 2, \ldots, n$, and is constant between possible values of *X*. We may also calculate the probability that *X* lies between two limits, l_1 and l_2 ($l_1 < l_2$); this is given by

$$\Pr(l_1 < X \le l_2) = \sum_{l_1 < x_i \le l_2} f(x_i) = F(l_2) - F(l_1) \tag{3.4}$$

i.e. it is the sum of all the probabilities for which x_i lies within the relevant interval.

Exercise 1. A bag contains six red balls and four blue balls. Three balls are drawn at random and not replaced. Find the probability function for the number of red balls drawn.

Let *X* be the number of red balls drawn. Then

$$\Pr(X=0)=f(0)=\frac{4}{10}\frac{3}{9}\frac{2}{8}=\frac{1}{30}$$

$$\Pr(X=1)=f(1)=\frac{4}{10}\frac{3}{9}\frac{6}{8}\times 3=\frac{3}{10}$$

$$\Pr(X=2)=f(2)=\frac{4}{10}\frac{6}{9}\frac{5}{8}\times 3=\frac{1}{2}$$

$$\Pr(X=3)=f(3)=\frac{6}{10}\frac{5}{9}\frac{4}{8}=\frac{1}{6}$$

It should be noted that $\sum_{i=0}^{3} f(i)=1$.

3.2 Continuous random variables

A random variable *X* is said to have a *continuous* distribution if *X* is defined for a continuous range of values between given limits (often $-\infty$ to ∞). An example of a continuous random variable is the height of a person drawn from a population, which can take any value (within limits!). We can define the *probability density function* (PDF) $f(x)$ of a continuous random variable *X* such that

$$\Pr(x<X\le x+dx)=f(x)dx \tag{3.5}$$

i.e. $f(x)dx$ is the probability that *X* lies in the interval $x<X\le x+dx$. Clearly $f(x)$ must be a real function that is everywhere ≥ 0.

If *X* can take only values between the limits l_1 and l_2 then, in order for the sum of the probabilities of all possible outcomes to be equal to unity, we require

$$\int_{l_1}^{l_2} f(x)dx=1 \tag{3.6}$$

Often *X* can take any value between $-\infty$ and ∞ and so

$$\int_{-\infty}^{\infty} f(x)dx=1 \tag{3.7}$$

The probability that *X* lies in the interval $a<X\le b$ is then given by

$$\Pr\left(a < X \le b\right) = \int_a^b f\left(x\right)dx \qquad (3.8)$$

i.e. $\Pr\left(a < X \le b\right)$ is equal to the area under the curve of $f\left(x\right)$ between these limits (see figure 3.2).

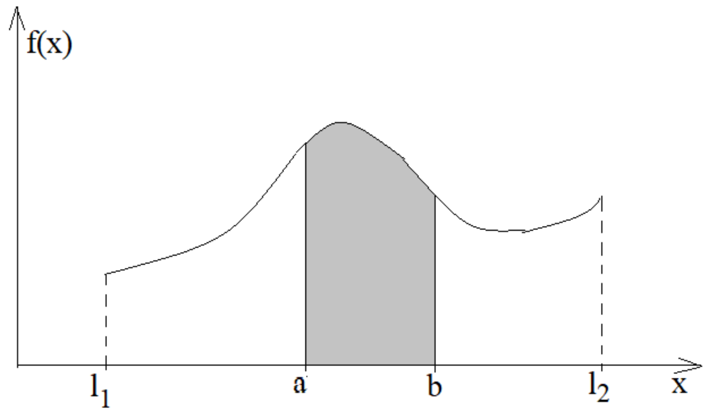

Figure 3.2 $\Pr\left(a < X \le b\right)$

We may also define the *cumulative probability function* $F\left(x\right)$ for a continuous random variable by

$$F\left(x\right) = \Pr\left(X \le x\right) = \int_{l_1}^{x} f\left(u\right)du \qquad (3.9)$$

where *u* is a (dummy) integration variable. We can then write

$$\Pr\left(a < X \le b\right) = F\left(b\right) - F\left(a\right) \qquad (3.10)$$

From (3.9) it is clear that $\dfrac{dF\left(x\right)}{dx} = f\left(x\right)$

Exercise 2. A random variable *X* has a PDF $f\left(x\right)$ given by Ae^{-x} in the interval $0 < x < \infty$ and zero elsewhere. Find the value of the constant *A* and hence calculate the probability that *X* lies in the interval $1 < X \le 2$.

We require the integral of $f\left(x\right)$ between 0 and ∞ to equal unity.

$$\int_0^\infty Ae^{-x}dx = -Ae^{-x}\Big|_0^\infty = A$$

and hence $A=1$. From (3.8), we then obtain

$$\Pr(1 < X \le 2) = \int_1^2 f(x)dx = \int_1^2 e^{-x}dx = e^{-1} - e^{-2} = 0.23$$

3.3 Sets of random variables

It is common in practice to consider two or more random variables simultaneously. For example, one might be interested in both the height and weight of a person drawn at random from a population. In the general case, these variables may depend on one another and are described by *joint probability density functions*. We note that if we have two random variables *X* and *Y* then we define their *joint probability density function* $f(x,y)$ in such a way that, if *X* and *Y* are discrete RVs,

$$\Pr(X = x_i, Y = y_j) = f(x_i, y_j), \tag{3.11}$$

or, if *X* and *Y* are continuous RVs,

$$\Pr(x < X \le x + dx\ ,\ y < Y \le y + dy) = f(x,y)dx\,dy \tag{3.12}$$

In many circumstances, however, random variables do not depend on one another, i.e. they are *independent*. As an example, for a person drawn at random from a population, we might expect height and IQ to be independent random variables. Let us suppose that *X* and *Y* are two random variables with probability density functions $g(x)$ and $h(y)$ respectively. In mathematical terms, *X* and *Y* are *independent* RVs if their joint probability density function is given by $f(x,y) = g(x)h(y)$. Thus, for independent RVs, if *X* and *Y* are both discrete then

$$\Pr(X = x_i, Y = y_j) = g(x_i)h(y_j) \tag{3.13}$$

or, if *X* and *Y* are both continuous, then

$$\Pr(x < X \le x + dx\ ,\ y < Y \le y + dy) = g(x)h(y)dx\,dy \tag{3.14}$$

The important point in each case is that the RHS is simply the product of the individual probability density functions (compare with the expression

$\Pr(A \cap B) = \Pr(A)\Pr(B)$ for statistically independent events *A* and *B*). The above discussion may be extended to any number of independent RVs X_i, $i = 1, 2, \ldots, N$.

Exercise 3. The independent random variables *X* and *Y* have the PDFs $g(x) = e^{-x}$ and $h(y) = 2e^{-2y}$ respectively. Calculate the probability that *X* lies in the interval $1 < X \le 2$ and *Y* lies in the interval $0 < Y \le 1$.

Since *X* and *Y* are independent RVs, the required probability is given by

$$\Pr(1 < X \le 2, 0 < Y \le 1) = \int_1^2 g(x)\,dx \int_0^1 h(y)\,dy =$$

$$= \int_1^2 e^{-x}dx \int_0^1 2e^{-2y}dy = \left(-e^{-x}\right)\Big|_1^2 \cdot \left(-e^{-2y}\right)\Big|_0^1 = \left(e^{-1} - e^{-2}\right)\cdot\left(1 - e^{-2}\right) = 0.2$$

3.4 Properties of distributions

For a single random variable *X*, the probability density function $f(x)$ contains all possible information about how the variable is distributed. However, for the purposes of comparison, it is useful to characterize $f(x)$ by certain of its properties. Most of these standard properties are defined in terms of *averages* or *expectation values*. In the most general case, the expectation value $E[g(X)]$ of any function $g(X)$ of the random variable *X* is defined as

$$E[g(X)] = \begin{cases} \sum_i g(x_i) f(x_i) & \text{for a discrete distribution} \\ \int g(x) f(x)\,dx & \text{for a continuous distribution} \end{cases} \tag{3.15}$$

where the sum or integral is over all allowed values of *X*. It is assumed that the series is absolutely convergent or that the integral exists. From its definition it is straightforward to show that the expectation value has the following properties:

(i) if *a* is a constant then $E(a) = a$;

(ii) if *a* is a constant then $E\left[ag(X)\right] = aE\left[g(X)\right]$;

(iii) if $g(X) = s(X) + t(X)$ then $E\left[g(X)\right] = E\left[s(X)\right] + E\left[t(X)\right]$

It should be noted that the expectation value is not a function of *X* but is instead a number that depends on the form of the probability density function $f(x)$ and the function $g(X)$. Most of the standard quantities used to characterize $f(x)$ are simply the expectation values of various functions of the random variable *X*. We now consider these standard quantities.

3.4.1 Mean

The property most commonly used to characterise a probability distribution is its *mean*, which is defined simply as the expectation value $E(X)$ of the variable *X* itself. Thus, the mean is given by

$$E\left[X\right] = \begin{cases} \sum_i x_i f(x_i) \text{ for a discrete distribution} \\ \int xf(x)\,dx \text{ for a continuous distribution} \end{cases} \qquad (3.16)$$

The alternative notations μ and $\langle x \rangle$ are also commonly used to denote the mean. If in (3.16) the series is not absolutely convergent, or the integral does not exist, we say that the distribution does not have a mean, but this is very rare in physical applications.

Exercise 4. The probability of finding a *1s* electron in a hydrogen atom in a given infinitesimal volume *dV* is $\psi^*\psi dV$, where the quantum mechanical wavefunction ψ is given by

$$\psi = Ae^{-r/a_0}$$

Find the value of the real constant *A* and thereby deduce the mean distance of the electron from the origin.

Let us consider the random variable *R* = ‘distance of the electron from the origin’. Since the *1s* orbital has no θ- or φ-dependence (it is spherically symmetric), we may consider the infinitesimal volume element dV as the spherical shell with inner radius r and outer radius $r+dr$. Thus, $dV = 4\pi r^2 dr$ and the PDF of R is simply

$$\Pr\left(r < R \le r+dr\right) = f\left(r\right)dr = 4\pi r^2 A^2 e^{-2r/a_0}dr$$

$$f\left(r\right) = 4\pi r^2 A^2 e^{-2r/a_0}$$

The value of A is found by requiring the total probability (i.e. the probability that the electron is somewhere) to be unity. Since R must lie between zero and infinity, we require that

$$A^2 \int_0^\infty e^{-2r/a_0} 4\pi r^2 dr = 1$$

Integrating by parts we find

$$A^2 4\pi \int_0^\infty r^2 e^{-2r/a_0} dr = 1$$

$$u = r^2 \qquad dv = e^{-\frac{2r}{a_0}}dr$$

$$du = 2rdr \qquad v = -\frac{a_0}{2}e^{-\frac{2r}{a_0}}$$

$$4\pi A^2\left(-\frac{a_0}{2}r^2 e^{-\frac{2r}{a_0}}\Bigg|_0^\infty + \int_0^\infty r a_0 e^{-\frac{2r}{a_0}}dr\right) = 1$$

$$u = r \qquad dv = e^{-\frac{2r}{a_0}}dr$$

$$du = dr \qquad v = -\frac{a_0}{2}e^{-\frac{2r}{a_0}}$$

$$4\pi A^2 a_0\left(-\frac{a_0}{2}re^{-\frac{2r}{a_0}}\Bigg|_0^\infty + \frac{a_0}{2}\int_0^\infty e^{-\frac{2r}{a_0}}dr\right) = 1$$

$$4\pi A^2 a_0 \frac{a_0}{2}\left(-\frac{a_0}{2}e^{-\frac{2r}{a_0}}\Bigg|_0^\infty\right) = 1$$

$$\pi A^2 a_0^3 (1-0) = 1 \quad \Rightarrow \qquad A = \sqrt{\frac{1}{\pi a_0^3}}$$

$$f(r) = 4\frac{r^2}{a_0^3} e^{-\frac{2r}{a_0}}$$

Now, using the definition of the mean (3.16), we find

$$E[R] = \int_0^\infty r f(r)\, dr = \int_0^\infty 4\frac{r^3}{a_0^3} e^{-\frac{2r}{a_0}} dr$$

$$E[R] = \frac{4}{a_0^3}\int_0^\infty r^3 e^{-\frac{2r}{a_0}} dr$$

$$u = r^3 \qquad\qquad dv = e^{-\frac{2r}{a_0}} dr$$

$$du = 3r^2 dr \qquad\qquad v = -\frac{a_0}{2} e^{-\frac{2r}{a_0}}$$

$$E[R] = \frac{4}{a_0^3}\left(-r^3 \frac{a_0}{2} e^{-\frac{2r}{a_0}} \Bigg|_0^\infty + \frac{3a_0}{2}\int_0^\infty r^2 e^{-\frac{2r}{a_0}} dr \right) = \frac{4}{a_0^3}\frac{3a_0}{2}\int_0^\infty r^2 e^{-\frac{2r}{a_0}} dr$$

$$E[R] = \frac{6}{a_0^2}\int_0^\infty r a_0 e^{-\frac{2r}{a_0}} dr = \frac{6}{a_0}\int_0^\infty r e^{-\frac{2r}{a_0}} dr = \frac{6}{a_0}\frac{a_0}{2}\int_0^\infty e^{-\frac{2r}{a_0}} dr$$

$$E[R] = 3\left(-\frac{a_0}{2} e^{-\frac{2r}{a_0}} \Bigg|_0^\infty \right) = -\frac{3a_0}{2}(0-1) = \frac{3a_0}{2}$$

3.4.2 Mode and median

Although the mean discussed in the last section is the most common measure of the ‘average’ of a distribution, two other measures, which do not rely on the concept of expectation values, are frequently encountered.

The *mode* of a distribution is the value of the random variable X at which the probability (density) function $f(x)$ has its greatest value. If there is more than one value of X for which this is true then each value may equally be called the mode of the distribution.

The *median* M of a distribution is the value of the random variable X at which the cumulative probability function $F(x)$ takes the value 1/2, i.e. $F(M)=\frac{1}{2}$. Related to the median are the *lower* and *upper quartiles* Q_l and Q_u of the PDF, which are defined such that

$$F(Q_l)=\frac{1}{4}\ , \qquad F(Q_u)=\frac{3}{4} \tag{3.17}$$

Thus the median and lower and upper quartiles divide the PDF into four regions each containing one quarter of the probability. Smaller subdivisions are also possible, e.g. the n-th *percentile*, P_n, of a PDF is defined by $F(P_n)=n/100$.

Exercise 5. Find the mode of the PDF for the distance from the origin of the electron whose wave function was given in the previous example.

We found in the previous example that the PDF for the electron's distance from the origin was given by

$$f(r)=4\frac{r^2}{a_0^3}e^{-\frac{2r}{a_0}} \tag{3.18}$$

Differentiating $f(r)$ with respect to r, we obtain

$$\frac{df}{dr}=\frac{4}{a_0^3}\left(2re^{-\frac{2r}{a_0}}-r^2\frac{2}{a_0}e^{-\frac{2r}{a_0}}\right)=\frac{8r}{a_0^3}\left(1-\frac{r}{a_0}\right)e^{-\frac{2r}{a_0}}$$

Thus $f(r)$ has turning points at $r=0$ and $r=a_0$, where $df/dr=0$. It is straightforward to show that $r=0$ is a minimum and $r=a_0$ is a maximum. Thus the mode of $f(r)$ occurs at $r=a_0$.

3.4.3 Variance and standard deviation

The *variance* of a distribution, $V[X]$, also written σ^2, is defined by

$$V[X] = E\left[(X-\mu)^2\right] = \begin{cases} \sum_j (x_j - \mu)^2 f(x_j) & \text{for a discrete distribution} \\ \int (x-\mu)^2 f(x)\,dx & \text{for a continuous distribution} \end{cases} \quad (3.19)$$

Here μ has been written for the expectation value $E[X]$ of *X*. From the definition (3.19) we may easily derive the following useful properties of $V[X]$. If *a* and *b* are constants then

(i) $V[a] = 0$

(ii) $V[aX + b] = a^2 V[X]$

The variance of a distribution is always positive; its positive square root is known as the *standard deviation* of the distribution and is often denoted by σ. Roughly speaking, σ measures the spread (about $x = \mu$) of the values that *X* can assume.

Exercise 6. Find the standard deviation of the PDF for the distance from the origin of the electron whose wavefunction was discussed in the previous two examples.

Inserting the expression (3.18) for the PDF $f(r)$ into (3.19), the variance of the random variable *R* is given by

$$V[R] = \int_0^\infty (r-\mu)^2\, 4\frac{r^2}{a_0^3} e^{-\frac{2r}{a_0}} dr = \frac{4}{a_0^3}\int_0^\infty \left(r^4 - 2r^3\mu + r^2\mu^2\right) e^{-\frac{2r}{a_0}} dr\,,$$

where $\mu = E[R] = 3a_0/2$. Integrating each term in the integrand by parts we obtain

$$V[R] = \frac{3}{4}a_0^2$$

Thus the standard deviation of the distribution is $\sigma = \sqrt{3}a_0/2$.

We may also use the definition (3.19) to derive the *Chebyshev inequality*, which provides a useful upper limit on the probability that random variable *X* takes values outside a given range centred on the mean. Let us consider the case of a continuous random variable, for which

$$\Pr\left(|X-\mu| \geq c\right) = \int_{|x-\mu|\geq c} f(x)\,dx$$

where the integral on the RHS extends over all values of x satisfying the inequality $|x-\mu| \geq c$. From (3.19), we find that

$$\sigma^2 = \int (x-\mu)^2 f(x)\,dx \geq \int_{|x-\mu|\geq c} (x-\mu)^2 f(x)\,dx \geq c^2 \int_{|x-\mu|\geq c} f(x)\,dx$$

$$= c^2 \Pr\left(|X-\mu| \geq c\right) \qquad (3.20)$$

The first inequality holds because both $(x-\mu)^2$ and $f(x)$ are non-negative for all x, and the second inequality holds because $(x-\mu)^2 \geq c^2$ over the range of integration. However, the RHS of (3.20) is simply equal to $c^2 \Pr\left(|X-\mu| \geq c\right)$, and thus we obtain the required inequality

$$\Pr\left(|X-\mu| \geq c\right) \leq \frac{\sigma^2}{c^2} \qquad (3.21)$$

A similar derivation may be carried through for the case of a discrete random variable. Thus, for any distribution $f(x)$ we have, for example,

$$\Pr\left(|X-\mu| \geq 2\sigma\right) \leq \frac{1}{4} \qquad \Pr\left(|X-\mu| \geq 3\sigma\right) \leq \frac{1}{9}$$

Exercise 7 Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50. If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be between 40 and 60?

Let X be the number of items that will be produced in a week:

By Chebyshev's inequality $P\left(|X-50| \geq 10\right) \leq \frac{\sigma^2}{10^2} = \frac{1}{4}$

Hence $P\left(|X-50| < 10\right) \geq 1 - \frac{1}{4} = \frac{3}{4}$

and so the probability that this week's production will be between 40 and 60 is at least 0.75.