BASIC RESULTS on PROBABILITY

Here we give results on probability that will be useful for Statistical Mechanics. Essentially this means we start from the theory for discrete events; this means that we will _rst need to recall basic facts about combinatorics (facts which are useful throughout physics). Later in the course we will go on to results on correlations, on continuous variables, and so on. It is assumed here that everyone has already taken a course on probability - this document is largely intended to refresh your memory.

1. Permutations and Combinatorics

We are interested in the theory of probability for a finite set of possible discrete outcomes. In physics this includes much of quantum mechanics, at least where a finite set of discrete states is involved. In statistical mechanics it involves the set of different microstates for any quantum system, which will be finite (but very large) if we deal with a finite system of N particles.

In more general areas of inquiry, the theory of discrete probabilities is involved in a huge range of activities and processes. Thus, the theory includes the calculation of the probability of any physical process which involves a finite set of outcomes. Well-known tutorial examples of this come from the calculation of probabilities for, eg., a possible outcome of a card game or indeed any game involving a finite set of possible states for the game.

1(a): COUNTING STATES: One begins by considering the way in which N objects can be permuted amongst each other, ie., in how many different ways these N objects may be arranged (eg., on a line). There are two obvious cases here, viz.,

(i) Distinguishable Objects: In this case the objects are all different or 'distinguishable' from each other. The number of different permutations is then simple to deduce. The object in the 1rst position on the line may be chosen in N different ways, that in the second position in N - 1 ways, and so on. The number of possible arrangements is therefore $N(N - 1)(N - 2)(1) = N!$

As an example let's look at 3 distinguishable objects, which we distinguish by their colour. Then, as shown in Fig. 1, there are $3! = 6$ different possible sequences for the 3 balls.

(ii) Indistinguishable Objects: Alternatively, the objects are all indistinguishable. But then all the different permutations just given are identical - so there is only one way of arranging N indistinguishable objects on a line (compare Fig. 1 again). There are actually 2 different ways that one can arrive at a consideration of N indistinguishable objects, viz.;

(a) we assume that the indistinguishability is absolute, ie., there is no difference in principle between the different permutations of the N objects (they really are not different from each other at all, and so the

$C^N_{n_1,n_2...n_6}$

only thing that is `physically real' is that there are N identical objects). This is what we have in quantum mechanics, for indistinguishable particles.

(b) we simply decide that there is no practically important distinction between different permutations (as in the case of a set of N coin tosses, where we don't care what order the coins are in).

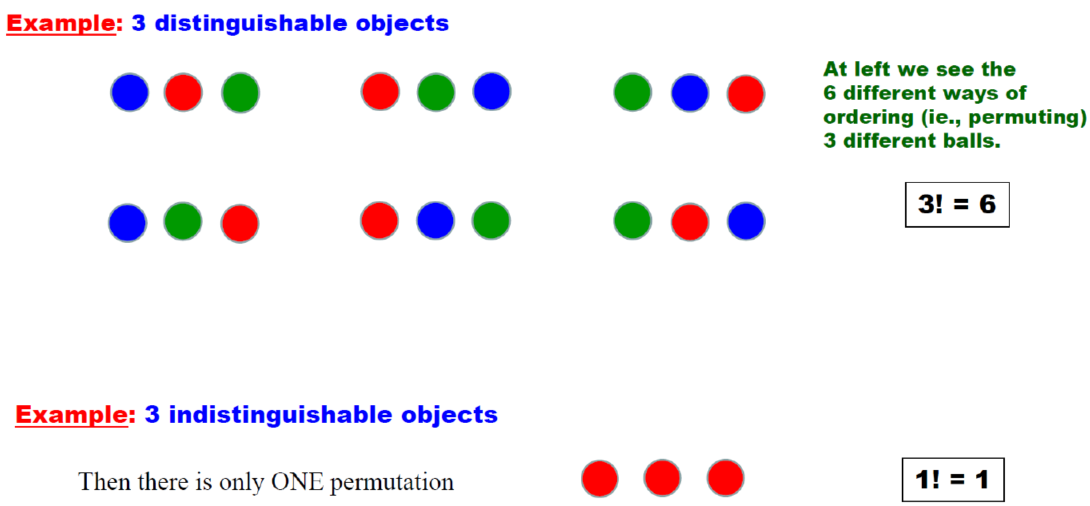


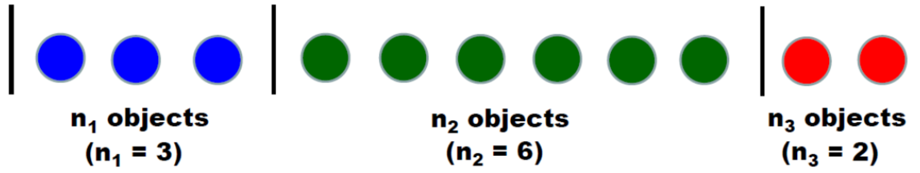
Fig 1: Illustration, with 3 objects, of the difference between distinguishable and indistinguishable states

1(b) MULTINOMIAL COMBINATORICS: Suppose we now consider N objects of which we have n_1 of type 1, which are identical to each other, n_2 identical objects of type 2, etc., etc., up to n_m identical objects of type m (so that $\Sigma_j n_j = N$). As we have seen, if all the N objects were distinguishable, the number of their permutations would be N! However, in the present case, the number of distinguishable permutations is only

$$C^N_{\{n_j\}} \equiv C^N_{n_1,n_2,..n_m} = \frac{N!}{n_1!n_2!....n_m!} \qquad (1)$$

since the j-th group of identical objects can be rearranged in n_j ! ways without changing anything, and we can do this for any of the m different sub-groups.

In Fig. 2 we show the example of a set of N dice, each of which can exist in one of 6 distinguishable states. Then the number of possible sequences with a throw of the N dice with n_1 dice coming up showing the number 1, n_2 dice showing the number 2, etc., is then given by the appropriate multinomial distribution (see Fig 2, Example 1).



Example 1: We throw N dice, each having six possible states (1-6) What is the total number of outcomes with n_1 showing 1, n_2 showing 2, etc, with $\Sigma_i\, n_i = N$?

The answer is $C^N_{n_1,n_2,..n_m}$ with m=6.

Example 2: We have 2 sets of indistinguishable objects; the total number is N, and n of them in one set, N-n in the other. What is the total number of distinguishable ways of ordering them? These could be, eg., H or T for N coins.

The answer is $C^N_n \;=\; \dfrac{N!}{n!(N-n)!}$

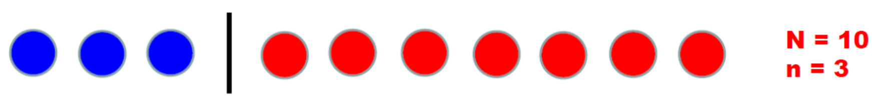


Fig. 2: the general idea of a multinomial distribution is shown in the upper picture, where we distinguish 3 types of object. In the middle picture (example 1) we apply this to a set of N dice, where one now has 6 indistinguishable states. In the lower picture (example 2) we depict a set of two indistinguishable states, where the binomial distribution applies

The simplest example of a multinomial distribution is of course the binomial distribution, viz.,

$$C^N_n \;=\; \frac{N!}{n!(N-n)!} \tag{2}$$

and the best-known example of this is when we look at a sequence of N coins, which are considered to exist in 2 distinguishable states (ie., heads H, or tails T). Then there are C^N_n different sequences of N coins in which we have n heads and N-n tails (or vice-versa, ie., N-n heads and n tails). Figure 2 shows the example of 10 coins, in which 3 are heads. I will return later to say a little more about the binomial distribution.

<u>Intuition for Multinomial Coefficients</u>: There are various intuitive ways to think about multinomial coefficients, since they arise in many situations. Here are three:

(i) Imagine we have N identical balls which we distribute in m different cells or boxes. What then is the total number of different ways that this can be done, with n_1 balls in the 1st cell, n_2 in the 2nd, and so on. The answer is the multinomial distribution in (1).

(ii) Suppose I have n_1 balls of 1 colour (all then indistinguishable from each other), n_2 balls of another, and so on up to the m-th colour. How many different distinguishable ways can we then order the N balls? The answer is again the multinomial distribution in (1).

(iii) There are clearly many different variants on these. Thus, eg., suppose we have N distinguishable rocks which we divide into m piles, such that we have n_k objects in the k-th pile, with $k = i = 1, 2, ..., m$. Suppose also that the ordering of rocks in each pile is irrelevant - they are just piles. Then it is clear that the multinomial coefficient gives the number of ways in which we can divide the different distinguishable objects into these piles. To see this explicitly, let us imagine we start by picking the n_1 rocks for the first pile; we can clearly do this in $C^N_{n_1}$ ways. The n_2 rocks in the second pile can then be picked from the $(N - n_1)$ remaining ones in $C^{N-n_1}_{n_2}$ ways; and so on until we get to the 2nd last pile, ie., the $(m-1)$-th pile, and this can be done in $C^{N-n_1-\cdots-n_{m-2}}_{N_{m-1}}$ ways. The remaining rocks, which make up the last pile, can only be found in one single way.

Multiplying all these together, we find that the total number of ways of dividing the original N rocks into m piles is given by

$$N \quad = \quad C^N_{n_1} C^{N-n_1}_{n_2} ... C^{N-n_1-\cdots-n_{m-2}}_{N_{m-1}} \quad = \quad \frac{N!}{n_1!n_2!....n_m!} \tag{3}$$

Here the last result is obtained by just multiplying out all the terms, and noting how the numerator in one such term is cancelled by a factor in the denominator of the previous term. One can of course think of many other ways to count different orderings that involve these coefficients.

The reason that 'multinomial' coefficients are given the name 'multinomial' is because they are just the coefficients in the 'multinomial' expansion of the expression $(z_1 + z_2 + \cdots + z_m)^N$, given by

$$\left(\sum_{k=1}^{m} z_k\right)^N \quad = \quad \sum_{n_1}\sum_{n_2} \cdots \sum_{n_m} \delta(N - \sum_k n_k) \left(\frac{N!}{n_1!n_2!n_m!}\right) z_1^{n_1} z_2^{n_2} \cdots z_m^{n_m} \tag{4}$$

Here the Kronecker delta function $\delta(N - \Sigma_k n_k)$ enforces the constraint that $\Sigma_k n_k = N$ (thus, eg., in our example above, the sum of the number of rocks in all the different piles must be equal to N).

Binomial Coefficients: Let's look a little more at the binomial coefficients, since these arise very often in physics and elsewhere. The easiest way to see how they arise is to consider some examples.

Example 1: The first example is the one that was depicted earlier, in Fig. 2, shown there as "example 2". Suppose we have 2 sets of indistinguishable objects (eg., one set of blue balls, and the other a set of red balls); the total number of objects is N, with n of them in one set, and $N - n$ in the other. The question we are asking is: what is the total number of distinguishable ways of ordering these objects? These could also be, eg., H (heads) or T (tails) for N coins.

One way to think about this is follows. We can organize the N objects in $N!$ different ways. However re-orderings of the $n!$ objects in the 1st set, and the $(N-n)!$ objects in the 2nd set, do not count, since all these different rearrangements are indistinguishable from each other. The number of possible arrangements is therefore just

$$C_n^N = \frac{N!}{n!(N-n)!} \tag{5}$$

ie., we get the binomial distribution.

Suppose, eg., I toss 10 coins, and I want to know how many different arrangements of these 10 tosses will have 3 heads and 7 tails turn up, in any order. The answer is then just $10!/(3!7!) = 120$.

Example 2: Suppose I have 7 pieces of glass, all of different colours (at one time one could find these on beaches, when people still used glass bottles). However 3 of the pieces are large pieces, and 4 of them are small. We want to know how many different possible ways we can order the 7 pieces in such a way that the large pieces are in one group and the small pieces in the other.

This is just an example of a general problem, in which we want to know the total number of ways of extracting n objects from N distinguishable objects, without regard for the order in which they are selected. Again, we can organize the N objects in $N!$ different ways. However re-orderings of the $n!$ objects in the group selected, and the $(N-n)!$ objects in the remaining group not selected, do not count, since they do not change this. The number of possible arrangements is therefore again C_n^N, the binomial distribution.

Later in the course we will study the binomial distribution in much more detail. We just note here that it is very sharply peaked around n ~ N/2.

2. Discrete Probabilities

Now that we have figured out how to count permutations for discrete outcomes (ie., where there is a finite number of possible outcomes for discrete events), we can move on to understand how one can calculate probabilities for the same discrete outcomes. To make the discussion simple, let's look first at bivariate probabilities, and then at multivariate probabilities.

2(a) BIVARIATE PROBABILITIES: First we look at problems in which there are only 2 different possible probabilities assigned to each of the discrete outcomes. Thus, eg., in a problem where we toss N coins, we can imagine that the coins are weighted so that all heads have one probability for turning up, and all tails another probability.

(i) <u>Equal Probabilities</u>: In the first and easiest kind of problem we will consider, all of the different possible outcomes are assumed to be a priori equally probable. Usually it is physically obvious for real systems when this should be the case, in that there is some symmetry which ensures that to suppose reason to suppose that one outcome is more likely than any other. A good example is the toss of a perfect coin, where by assumption the probability of getting heads (H) or tails (T) is equal (and therefore each has probability 1/2).

In this case there are 4 possible outcomes if we toss two such coins, and/or if we toss the same coin twice. These outcomes are just HH, HT, TH, and TT. Clearly each of these is equally likely, and so their probabilities are 1/4 each.

It is then clear that if we want to calculate the probability of one specific outcome when the total number of possible outcomes is Q, that probability must be 1/Q. On the other hand, if we want to find the probability that we will get an outcome which itself involves a set of S different discrete outcomes (this set being a subset of all possible outcomes), then that probability will be S/Q.

This is where all the counting exercises we have done above come in handy. To see this, consider the probability that we will get n heads if we throw a perfect coin N times. The ordering is irrelevant, so that the total number of ways of throwing n heads is just $C^N{}_n$, as we have already seen. However the total number of possible outcomes is clearly 2^N; and so it follows that the probability $P_N(n)$ of getting n heads is just given by

$$P_N(n) \quad = \quad \frac{C_n^N}{2^N} \quad = \quad \frac{1}{2^N}\left(\frac{N!}{n!(N-n)!}\right) \tag{6}$$

which is very sharply peaked around n = N=2 if N is large. When N; n _ 1 we can easily _nd accurate expressions for this using Stirling's asymptotic formula (seelater in the course).

(ii) <u>Unequal Probabilities</u>: Now suppose that the probabilities for the different outcomes are not the same. At this point we must assign probabilities depending on what knowledge we have of the system involved. Suppose, eg., that for the coin discussed above, we know somehow that the probability of getting heads is p+, so that the probability of getting tails is p_{-} = (1-p+). We count things in the same way, but now we have to assign the correct probability to each outcome. It should be immediately obvious that the new result for $P_N(n)$ is now

$$P_N(n) \quad = \quad C_n^N \, p_+^n p_-^{N-n)} \quad \equiv \quad \left(\frac{N!}{n!(N-n)!}\right) p_+^n (1-p_+)^{(N-n)} \tag{7}$$

because the probability of getting any one of the combinations with n heads and N-n tails is, by assumption, just $p^n{}_+ \, p\_^{N-n}$. When p+ = 1/2, this just reduces to the previous result.

2(b) MULTIVARIATE PROBABILITIES: It should now be obvious how to generalize this to cases where we have more than two different types of object involved in our outcomes, ie., where we must deal with multinomial combinatorics. Thus, suppose we have N identical balls which we distribute in m different cells or boxes, but now the probability of going into the k-th box is p_k, where k = 1,… m (and where of course $\Sigma_k p_k = 1$).

As we saw before, the number of different ways of doing this is just the multinomial coefficient; but now the weighting attached to any one of these ways is $\Pi_k (p_k)^{n_k}$. It then follows that the probability $P_N(n_1; n_2; .. n_m)$ of getting an outcome in which there are n_k balls in the k-th box is just

$$\begin{aligned} P_N(n_1, n_2, \cdots n_m) &= C^N_{n_1,n_2,..n_m} \prod_{k=1}^{m} p_k^{n_k} \\ &= \delta(N - \sum_k n_k) \left(\frac{N!}{n_1! n_2! n_m!} \right) p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m} \end{aligned} \tag{8}$$

Again, we include the Kronecker delta constraint, as before. Note that if we now sum over all possible outcomes here (which means summing over all the different values of the n_k within the constraint that $\Sigma_k n_k = N$), then we just get back the formula (3) above, with p_k substituted for z_k. Note also that the left-hand side of (3) then becomes unity, because $\Sigma_k p_k = 1$, and this is of course what we would expect - the sum of the probabilities of all different outcomes exhausts all possibilities, and so it must be unity.

Multivariate Examples: Let's consider some examples of what we are talking about here, to give you an idea. I will, for simplicity, look here at cases where the probabilities of each outcome are all the same; the way to generalize will be fairly clear. The main intellectual exercise here is to deduce what are the appropriate multinomial distributions to describe the situation (see Fig. 3 below for a visual representation of these examples).

example 1: Suppose I draw 7 cards from a 52-card pack of cards. What is the probability that this hand of cards contains 3 Aces?

To do this we need to first ask how many possible outcomes there are for the 7 cards that are dealt; we then ask how many of these give 3 Aces. The probability is then the latter number divided by the former. The first question is simple - the total number of possible distinguishable arrangements is the binomial $C^{52}_7 \equiv C^{52}_{45} = 52!/7!\,45!$, because we can re-order the first 7 cards 7! times, and the last 45 cards 45! times.

To deal with the 2nd question we note first that it does not matter which Aces we get. We need to multiply the number of ways of getting 3 of the 4 Aces (without caring which ones), by the total number of outcomes for the other 4 cards that are dealt, with the constraint that these are NOT Aces. The first number is $C^4_3 = 4$. To find the second number, we note that there are 48 cards that are not Aces, and we are getting 4 of these. So this latter number is $C^{48}_4 = 48!/44!\,4!$.

The final result for the probability $P^{\{7\}}_{AAA}$ is then

$$P^{\{7\}}_{AAA} = \frac{C^4_3 C^{48}_4}{C^{52}_7} = 4 \times \frac{48!}{4!\,44!} \times \frac{7!\,45!}{52!} = 7.6.5.4 \frac{45}{52.51.50.49} \tag{9}$$

which if we work it out gives $P^{\{7\}}_{AAA} \sim 0.00582$, ie., roughly a probability of $1/172$.

Example 2: In a game of poker, each of four players is dealt 5 cards from a pack of 52 cards. What is the probability that each player is dealt an ace?

A: This is a generalization of the last problem to a multinomial distribution. We must first ask how many possible outcomes there are for the 4 batches of 5 cards that are dealt; we then ask how many of these give 1 Ace in each hand. The probability is then the latter number divided by the former.

The answer to the first question is given by the multinomial distribution - we have $C^{52}_{5.5.5.5.32} \equiv 52!/(5!)^4 32!$ ways of distributing the cards amongst 4 hands of 5 cards, and amongst the remaining 32 cards.

To deal with the second question we note first that the ordering of the Aces between the hands is irrelevant - there are 4! different ways of ordering the 4 Aces. There are then 48 cards left, that are not Aces - these can be dealt out to the 4 different hands in a total of $C^{48}_{4.4.4.4.32} \equiv 48!/(4!)^4 32!$ times.

The final result for the probability $P^{\{4\times5\}}_{4A}$ is then given by

$$P^{\{4\times5\}}_{4A} = \frac{4! \times C^{48}_{4.4.4.4.32}}{C^{52}_{5.5.5.5.32}} = \frac{5^4 \times 24}{52.51.50.49} \tag{10}$$

which, if we work it out, just gives $P^{\{4\times5\}}_{4A} \sim 2.31 \times 10^{-3}$, ie., a probability $\sim 1/433$.

Example 3: There is a British game called snooker, one of a large variety of different games of billiards. In this game, one has a white, a yellow, a green, a brown, a blue, a pink, a black and 15 red balls (for a total of 22 balls). So - how many different permutations can one make of these (ie., how many different distinguishable ways can one order them, assuming all the reds are **INdistinguishable** from each other)? And, if we pull balls out at random, what is the probability that the first 15 of these will be red?

A. We are clearly dealing with the multinomial distribution here, with the result that we have

$$\frac{22!}{(1!)^7 15!} \tag{11}$$

different distinguishable permutations. You can easily do the second part of this question yourself.

This completes our summary of the way in which probabilities are assigned for a finite number of discrete outcomes, ie., of the topics of discrete combinatorics and discrete probabilities. The examples 1-3 are depicted on the next page in Figure 3.

I have of course only scratched the surface – if you want to know more you can use well-known textbooks, or look at Wikipedia articles such as

https://en.wikipedia.org/wiki/Combinatorics
https://en.wikipedia.org/wiki/Probabilistic_method
https://en.wikipedia.org/wiki/Binomial_distribution
https://en.wikipedia.org/wiki/Multinomial_distribution

Even though we have only scratched the surface, you can, armed with this understanding, not only understand the probabilities assigned in most of statistical mechanics, but also successfully play poker!

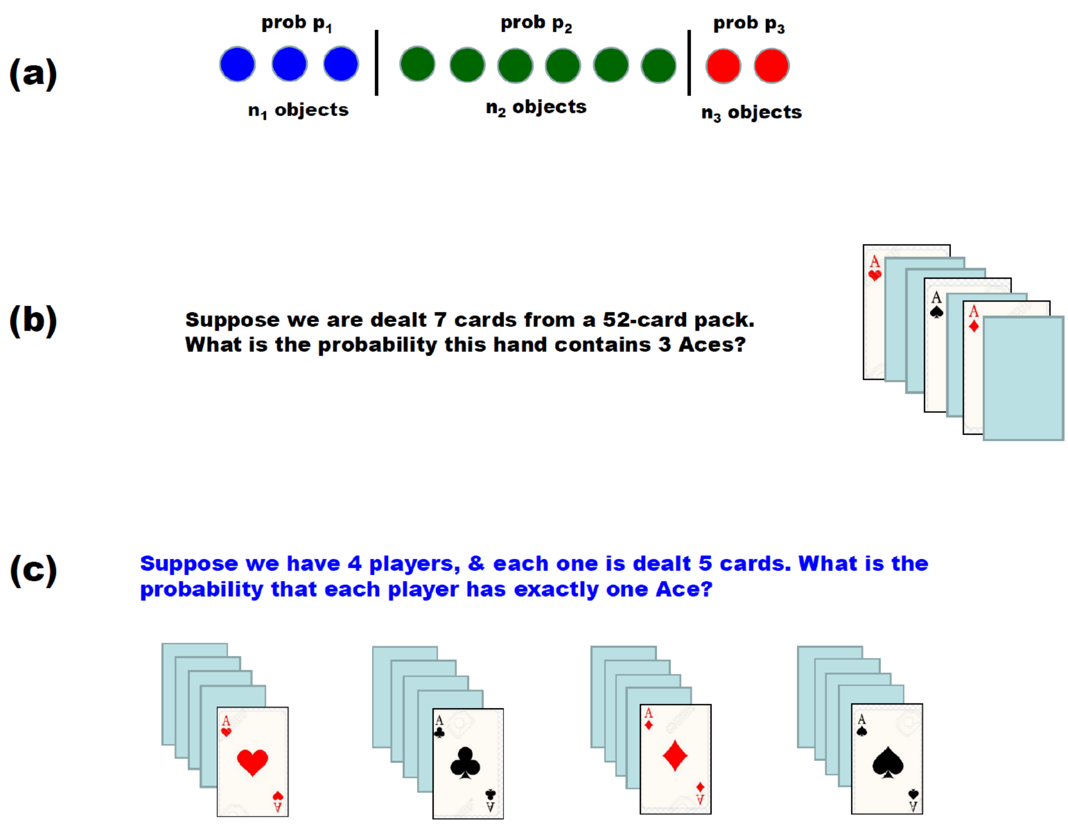


Figure 3: In (a) we see how the multinomial distribution is used for a problem involving 3 different probabilities p_1; p_2; p_3. In (b) we ask the probability that we get 3 Aces in a draw of 7 cards from a 52-card pack. In (c) we ask what is the probability that if 4 players each draw 5 cards, that each of them will get one Ace.

------------------- FINIS -------------------