LECTURE NOTES IN APPLIED PROBABILITY

0. Counting

The purpose of counting in probability theory is to determine the number of ways an experiment can turn out. For example, in a 42-number lottery, there are 5,245,786 ways for the State to pick the six winning numbers from a hat containing 42 numbers. We shall prove this.

If you think about it, the State draws little plastic markers *labelled* with numbers, not numbers. That is, the State uses integers to code the objects of the experiment. This is a matter of convenience. Even football teams use numbers on jerseys to code their human participants. As we shall see, the integer coding of objects in an experiment simplifies the description. It will not matter whether the objects are people or playing cards. Because there is a game to illustrate every counting principle, we can study the classical counting formulas by studying games.

0.1. Sampling With and Without Replacement

Consider the drawing of numbers from a hat containing the numbers 1 through n. The first draw produces a number. This number is replaced—i.e. tossed back into the hat, and a second number is drawn. This number is replaced, and so on, until r numbers are drawn. For each draw, there are n numbers that can be drawn, so the number of possible outcomes for the experiment is

$$N = n^r. \tag{1}$$

We call this experiment "sampling with replacement."

Example 0.1: Birthdays. In a waiting line of r people, there are n^r possible sequences of birthdays, with $n = 365$.

Now suppose r numbers are drawn from a hat *without* replacement. The first draw produces one of n possible outcomes, the second produces one of $n-1$ possible outcomes, and so on. The number of possible outcomes for r such draws is

$$N = (n)(n-1)\cdots(n-r+1). \tag{2}$$

We call this experiment "samping without replacement," and we call the number of possible outcomes "r

permutations of n." For convenience, we will use the following notation for this number:

$$(n)_r = (n)(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}. \tag{3}$$

Note that "n permutations of n" is "n factorial:"

$$(n)_n = n!. \tag{4}$$

Example 0.2: Birthdays. In a room full of r people, there are $(n)_r$ possible ways to produce r *different* birthdays, with $n = 365$. Therefore, there are $n^r - (n)_r$ ways that at least one birthdate is shared by two of the r people.

Example 0.3: Horse Race. In a horse race with n horses, there are $(n)_r$ ways to determine the order of finish for the first r horses and $n!$ ways to determine the exact order of finish for the field. If $n = 12$ and $r = 6$, then there are $(12)_6 = 665,280$ ways to determine the exact order of finish for the first six horses. This number is so large that no one should ever bet on it. If $n = 12$ and $r = 3$, then $(12)_3 = 1,320$. This number is large but not too large, so some people bet on the exact order of finish for the first three horses. This is called a "trifecta perfecta" bet.

Sampling with replacement and *sampling without replacement* are actually *composite* experiments. In the first, the composite experiment consists of *repeated trials*, and in the second it consists of *modified trials*. The number of possible outcomes is just the product of the possible outcomes for each trial:

$$n = \prod_{i=1}^{r} n_i. \tag{5}$$

When the trials are repeated, as in sampling with replacement, then $n_i = n$. When they are modified, then n_i is defined by the experiment. If the experiment is sampling without replacement, then

$$n_i = n - i + 1 = n_{i-1} - 1. \tag{6}$$

Example 0.4: Reordering Books. Three border collie books, four Scottish golf books, and five probability books are to be arranged on a shelf. There are $(3+4+5)! = 12!$ ways to reorder them. If, however, the books of a type must stay together, there are $3!4!5!3!$ ways to reorder them, with the first three terms accounting for the number of ways to order within a type and the last 3! corresponding to the number of ways to order the types. The reordering is a composite experiment.

0.2. Permutations and Horse Races

Consider a twelve-horse Kentucky Derby. There are

$$12(11)\cdots(2)(1) = 12! = 479,001,600$$

ways for the horses to finish. You would never bet on the exact order of finish because your chances of choosing the correct 1 out of 479,001,600 are too remote, but you might bet on Seabiscuit to win. There are twelve ways for the race to produce a winner, and this experiment is no different than drawing one number from a hat that contains twelve numbers. We could, however, reason somewhat differently: there are 12! possible outcomes for the race. For each such outcome, the reordering of the eleven "also-rans" leaves the winner unchanged, so the number of ways to determine the winner is $12!/11! = 12$. (There is nothing special about the winner in this type of argument. In fact, there are 11! reorderings of every position other than the seventh, for example, meaning that there are also twelve ways to pick the seventh-place horse.)

We may extend this line of reasoning to determine the number of ways that the Derby can produce win, place, and show. There are 12! possible outcomes for the race. For each such outcome, the reordering of the nine also-rans leaves the order of finish for the top three unchanged, so the number of ways to produce win, place, and show is $12!/9! = 1320 = (12)_3$.

Suppose we favor Seabiscuit—but not enough to lay down a bet on Seabiscuit to win. We would like to hedge our bet by betting on win, place, or show. There are 11! ways for Seabiscuit to win, 11! other ways for Seabiscuit to place, and 11! other ways for Seabiscuit to show. Thus, there are 3(11!) outcomes that produce Seabiscuit as win, place, or show, and this statement could be made about any horse in the field.

There is a simplifying principle here. We have treated also-rans as indistinguishable horses whose exact order of finish has no influence on the outcome of interest—namely, the exact order of finish of the first three horses. This principle of dividing out the order of finish for indistinguishable objects may be applied more generally. For example, if you are given the word "geranium," you know that there are $8! = 40,320$ ways to reorder the letters to construct 8! other "words," but if you are given the word "Mississippi," then you know that there are $(11!)/(4!4!2!) = 34,650$ ways to reorder the letters to construct other words. The denominator corrects for indistinguishable orderings of the letters i, s, and p that have respective multiplicities of 4, 4, and 2.

0.3. Elections and Combinations

The slate of candidates for a school board contains seven names. The top three vote-getters will be elected to the Board, and the order of finish will determine President, Vice President, and Secretary. The number of ways for the election to order candidates is $7! = 5,040$. The number of ways to determine President, Vice President, and Secretary is $7!/4! = 210$, because the order of finish for the four also-rans is immaterial to the order of finish for the top three. But how many ways are there to determine membership on the Board? When we ask this question, we are asking how many three-person Boards can be made out of seven people. The order of finish for the winners does not matter: all 3! reorderings of the winners leave the membership on the Board unchanged. Therefore, we divide the number $7!/4!$ by 3! to account for the 3! reorderings of each winning threesome that leaves membership unchanged. In summary, the number of three-person School Boards that can be constructed from seven candidates is

$$N = \frac{7!}{3!4!} = 35.$$

We call this the "number of combinations of seven things taken three at a time" or "the number of ways to draw three objects from 7" and denote it $\binom{7}{3}$. The general formula is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n)_k}{k!}, \tag{7}$$

which is read "n choose k." It is the number of subsets of size k within a set of size n.

There is no (mathematical) difference between the problems of producing three-person committees from a pool of seven, three girls in a family of seven children, or three ones in a seven-bit sequence. There are $\binom{7}{3}$ ways to do it. One nice way to illustrate this is to code committee members, girls, or binary ones with 1's and then to draw Pascal's triangle. The triangle of Figure 0.1 may be used to determine $\binom{n}{k}$ for $0 \leq k \leq n$, $n = 1, 2, 3, 4$. If we look at panel (c) of the figure, we see that each interior number is the sum of the two above it. This is a property of the binomial coefficients; namely,

$$\binom{n+1}{m+1} = \binom{n}{m} + \binom{n}{m+1}.$$

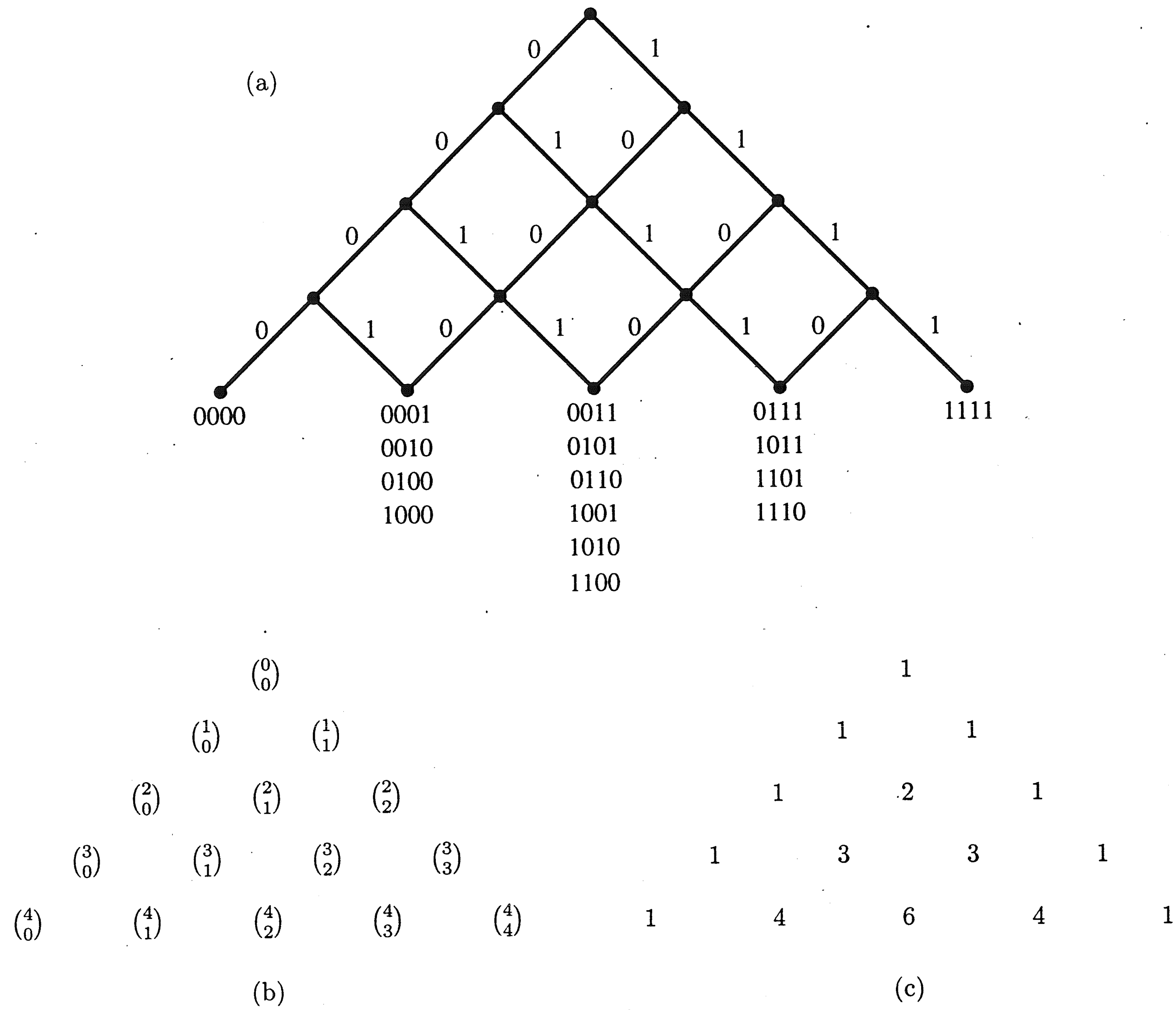

Figure 0.1. Panel (a) depicts the $2^4 = 16$ binary words of length 4, arranged according to the number of ones in each word. For each such word, there is a unique path from the top of the tree to the leaf. In panel (b), we have arranged the binomial coefficients in a similar array, called Pascal's triangle. The coefficient $\binom{n}{m}$ is the number of binary words of length n containing m ones. Panel (c) contains the values of the binomial coefficients. Each number is the sum of the two above (with zeros outside the triangle).

Example 0.5: Playing Cards. There are $\binom{52}{5} = 2,598,960$ ways to choose five cards from a 52-card deck.

Example 0.6: Coin Tosses and Binary Transmissions. There are 2^n different binary transmissions in an n-bit transmission. Of these, $\binom{n}{n/2}$ of them have exactly $n/2$ ones and $n/2$ zeros when n is even.

Example 0.7. A "box bet" at the racetrack is a bet on the first three horses, independent of the respective order of these three. The number of ways to hit a box bet is $\binom{n}{3}$ when n is the number of horses in the field.

0.4. The Multinomial Coefficients

These arguments generalize to the construction of r committees of sizes $n_1, n_2, \ldots, n_r$ from a pool of n, with $n_1 + n_2 + \cdots + n_r = n$. Simply rearrange the pool $n!$ ways, let the first n_1 be members of committee 1, and so on. Divide out the orderings of each committee that leave the composition unchanged. Then the number of ways to build r committees of sizes $n_1, n_2, \ldots, n_r$ from a pool of n is

$$N = \binom{n}{n_1\, n_2 \cdots n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}; \qquad \text{where } n_1 + n_2 + \cdots + n_r = n. \tag{8}$$

This is called the *multinomial coefficient,* a generalization of the binomial coefficient. That is, when $r = 2$,

$$\binom{n}{k \quad n-k} = \binom{n}{k}. \tag{9}$$

0.5. The Games People Play

In this section, we will apply our counting formulas to some common games people play. This will lead to the concept of a *fair* game. It will illustrate that games like Lotto and Keno generate money for the State only because they are *unfair* games designed to extract money from the people. Some think this is even more sinister than taxation without representation.

Lotto. Here is the way Lotto is played in most states. You select six numbers from 42. You may choose your own or have a computer program randomly select them for you. The State then stages a flamboyant public drawing of six numbers from a rotating cage of 42 numbers. You hit it big by matching all six numbers. You hit it small by matching fewer numbers. The question is this: what are your odds of hitting it big or of hitting it small? That is, what are your odds of matching k of the six good numbers, where k is a number from 0 to 6?

We shall answer this question by first noting that there are $\binom{42}{6} = 5,245,786$ different six-number draws that can be made by the State. You must hit the one draw that is made, so your odds are 5,245,786 to 1. But what are your odds of hitting, say, four of the good numbers? Let's break the 42 numbers into two sets, the set of six "good" numbers drawn by the State and the corresponding set of 36 "bad" numbers (the also-rans, if you like). If you match four good numbers, then you also match two of the bad numbers. The number of ways of doing this is $\binom{6}{4}\binom{36}{2}$. The first term is just the number of ways to choose four numbers from six, and the second term is just the number of ways to choose two numbers from 36. Generally, the number of ways to match k numbers in a $(42, 6)$ Lotto is

$$N = \binom{6}{k}\binom{42-6}{6-k}.$$

The first term is the number of ways to select k of six numbers, and the second term is the number of ways to select $6-k$ of 36 numbers. The odds of matching k numbers is just the number of possible outcomes vs. the number of outcomes that produce k matches:

$$\binom{42}{6} \text{ to } \binom{6}{k}\binom{42-6}{6-k}.$$

The "inverse of odds" is the *probability* of k matches:

$$P(k; 42, 6) = \frac{\binom{6}{k}\binom{42-6}{6-k}}{\binom{42}{6}}; \qquad k = 0, \ldots, 6.$$

Keno. In Keno, the player selects ten numbers from sixty. The State then draws twenty numbers from a rotating cage of sixty numbers. You hit it big by matching ten of the State's twenty numbers, and you hit it small by matching fewer. The number of ways for you to draw ten numbers from sixty is $\binom{60}{10}$. The number of ways for you to match k of twenty good numbers is $\binom{20}{k}$, and the number of ways to match $10-k$ of forty bad numbers is $\binom{40}{10-k}$. So, the probability of matching k numbers in a $(60, 20, 10)$ Keno game is

$$P(k; 60, 20, 10) = \frac{\binom{20}{k}\binom{60-20}{10-k}}{\binom{60}{10}}; \qquad k = 0, \ldots, 10.$$

Polling. This principle applies to more general settings. Suppose a population of n students consists of n_1 male engineering students, n_2 female engineering students, n_3 male liberal arts students, and n_4 female liberal arts students. Then the probability that k draws produces the numbers k_1, k_2, k_3, k_4 from these types is

$$P(k_1, k_2, k_3, k_4) = \frac{\binom{n_1}{k_1}\binom{n_2}{k_2}\binom{n_3}{k_3}\binom{n_4}{k_4}}{\binom{n}{k}}; \qquad n = n_1 + \cdots + n_4, \qquad k = k_1 + \cdots + k_4. \tag{10}$$

You can see two principles at work: *combination* and *composition of experiments.*

Poker. Poker hands are five-card hands dealt from a 52-card deck consisting of four suits (hearts, diamonds, clubs, and spades) and thirteen faces (2 through 9, jack, queen, king, ace) in each suit. We shall denote a hand by xyzwv, where the letters stand for one of thirteen faces. If a hand looks like xxyzw, then it has one pair (the xx), and these paired faces must of course be of different suits.

The number of five-card poker hands is

$$N = \binom{52}{5} = 2,598,960.$$

This is just the number of ways to choose five cards from 52. Among these, some are special:

(i) no pairs: xyzwv

(ii) one pair: xxyzw

(iii) two pairs: xxyyz

(iv) three of a kind: xxxyz

(v) full house: xxxyy

(vi) four of a kind: xxxxy.

These hands exhaust all possibilities, so the total number of such hands must equal the number of five-card hands. You will check this in problem 0.7. Among the no-pair hands, a few are special:

(vii) flush: xyzwv, all of the same suit

(viii) straight flush: xyzwv, all of the same suit and consecutive in number

(ix) royal flush: xyzwv, all of the same same suit, 10JQKA.

Our objective is to count the number of ways that each such hand can be generated and, from these counts, determine the odds of getting such a hand. Our principle for counting will be this:

1. Count the number of ways to select the faces for the hand.
2. Count the number of ways that the various roles may be played in the hand (e.g., how many ways there are to decide which face will play the role of the pair).
3. Count the number of ways to select the suits for the hand.

We shall illustrate the application of this principle by counting a few hands and leave the rest of the computations to problem 0.7.

The number of ways to get a no-pair hand (xyzwv) is

$$N = \binom{13}{5}\binom{5}{0}\binom{4}{1}\binom{4}{1}\binom{4}{1}\binom{4}{1}\binom{4}{1}.$$

The term $\binom{13}{5}$ counts the number of ways to select five *different* faces from thirteen. The term $\binom{5}{0}$ counts the number of ways for these five faces to play zero special roles, and the five terms $\binom{4}{1}$ count the number of ways to choose the suit for each of the five faces.

The number of ways to get a two-pair hand (xxyyz) is

$$N = \binom{13}{3}\binom{3}{1}\binom{4}{2}\binom{4}{2}\binom{4}{1}.$$

The term $\binom{13}{3}$ accounts for the number of ways to get three different faces from thirteen. The term $\binom{3}{1}$ accounts for the number of ways to decide which of the three faces will play the role of the singleton. The term $\binom{4}{2}$ accounts for the number of ways to choose two suits from four to assign to the pairs, and the term $\binom{4}{1}$ accounts for the number of ways to assign one suit to the singleton.

Fair Games, Odds, and Payout. In most games like Lotto and Keno, you pay B dollars for your bet or your play. If you lose, you forfeit your B-dollar bet. If you win, you win W dollars for a net win of $W - B$. For the game to be fair, your winnings should equal your losings. ("You win some, and you lose some.") Total winnings are $(W - B)$ times the number of ways to win, and total losses are B times the number of ways to lose. The number of ways to lose is just the number of outcomes for the game minus the number of ways to win, so we may say that a game is fair when

$$(W - B)(\text{number of ways to win}) = B(\text{number of outcomes} - \text{number of ways to win}) \tag{11}$$

or

$$\frac{W}{B} = \frac{\text{number of outcomes}}{\text{number of ways to win}} = \text{odds}. \tag{12}$$

That is, *the ratio of the payout to the bet in a fair game is odds.* In fact, most games are played at "odds offered." They should equal the odds for the game, but they never do for any game you will ever play against a casino or the State.

Example 0.8: Keno. The odds of matching eight numbers in a (60, 20, 10) Keno are 768 to 1. (See problem 0.5.) Therefore, a \$1 Keno ticket should pay out \$768 for eight matches. It actually pays \$100. This is decidedly unfair. Thank you very much for playing.

0.6. Myths and Amusements

In this section we will look at some fascinating little questions that allow us to use the insights we have developed about counting and probability to find simple yet sophisticated answers to vexing questions.

Myth 1: Draw First. You are drawing straws, hoping to get the long one and gain a favor (or, equivalently, hoping to miss the short one and avoid a consequence). The myth is that you should draw first for the long one and last for the short one. If there are n straws and you draw first, you get the long straw with probability $1/n$. If you draw second, you get the long straw with probability

$$P_2 = \frac{n-1}{n}\,\frac{1}{n-1} = \frac{1}{n}, \tag{13}$$

where $\frac{n-1}{n}$ is the probability that the long straw remains and $\frac{1}{n-1}$ is the probability that you get it. For the third draw, the probability of getting the long straw is

$$P_e = \frac{n-1}{n}\,\frac{n-2}{n-1}\,\frac{1}{n-2} = \frac{1}{n}, \tag{14}$$

where $\frac{n-1}{n}\,\frac{n-2}{n-1}$ is the probability that the long straw remains and $\frac{1}{n-2}$ is the probability of getting it. So, the probability of winning (or losing) is independent of the order of the draw.

Myth 2: Stand Pat. You are a contestant on a game show. Monte Hall presents you with three curtains, two of which hide goats and one of which hides a Mazda RX-7. First, you select a curtain. Monte Hall then opens one of the other two curtains to reveal a goat. He gives you the chance to switch curtains. You know that Monte Hall is a smooth-talking game show host who wants to get rid of the goats and keep the RX-7 for himself, so you stand pat. This would be a mistake. If you stand pat, your probability of winning is $1/3$, namely the probability that your pick was correct. If you switch, your probability of losing is $1/3$, namely the probability that your original pick was correct. So, if you switch, your probability of winning is $1 - \frac{1}{3} = \frac{2}{3}$. Do not stand pat. Switch.

Birthday Matches. What is the probability of having at least one common birthday in a room containing r people? There are n^r ways to choose birthdays but just $(n)_r$ ways to choose unique birthdays. This is the difference between sampling with and without replacement. The probability of *no* matches is the ratio of these two:

$$P_1(0) = \frac{(n)_r}{n^r}. \tag{15}$$

This probability is about 1/2 for $r = 23$. The probability of one or more matches is

$$1 - P_1(0) = 1 - \frac{(n)_r}{n^r} = 1 - \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{r-1}{n}\right). \tag{16}$$

This may be approximated by using the approximation $e^{-k/n} \cong 1 - \frac{k}{n}$ for large n:

$$P_1(0) \cong e^{-1/n} e^{-2/n} \cdots e^{-\frac{r-1}{n}} = e^{-\frac{1}{n}\frac{r(r-1)}{2}} \tag{17}$$

$$1 - P_1(0) \cong 1 - (e^{-\frac{r-1}{n}})^{r/2}.$$

Let's make this question a little more interesting by asking what the probability is that we find one or more matches for *your* birthday. The probability is

$$1 - P_2(0) = 1 - \left(1 - \frac{1}{n}\right)^{r-1}, \tag{18}$$

where $1 - \frac{1}{n}$ is the probability of no match by a given person and $P_2(0) = \left(1 - \frac{1}{n}\right)^{r-1}$ is the probability that $r - 1$ people mismatch your birthday. The answer is approximately

$$1 - P_2(0) \cong 1 - e^{-\frac{r-1}{n}}. \tag{19}$$

The probability of one or more birthday matches is 1/2 at $r = 23$, but the probability of one or more matches to *your* birthday is not 1/2 until $r = 254$. Why is this? Roughly, in a room of r_2 people, there are just $r_2 - 1$ tries to get a personal match, and in a room of r_1 people there are $\frac{r_1(r_1-1)}{2}$ tries to get any match. And,

$$r_2 - 1 = \frac{r_1(r_1 - 1)}{2} \quad \text{at} \quad r_1 = 23, \; r_2 = 254.$$

0.7. Occupancy Problems and the Distributions of Statistical Mechanics

A monoatomic gas is a collection of molecules in action. Each molecule traces out a trajectory in a phase space of "position" and "momentum." At any point in time, the molecules are distributed unevenly over the cells of the phase space. Thermodynamics tells us that we can only measure the effects of this distribution on properties like temperature. Such properties do not, evidently, depend on position or on molecular identity. They do depend on the distribution of molecules over energy cells. So, the problem of statistical thermodynamics is to set up a model for phase space and to determine the number of ways that molecules can get into energy cells. We have to be very careful about associating probability with these distributions, because we are saying nothing about the equal probability of any particular arrangement.

Figure 0.2 illustrates a phase space with r energy cells and k position cells. We want to determine how many ways there are to distribute n molecules into these kr position-energy cells in such a way that the number of molecules in energy cell i is n_i. We shall call $\mathbf{n} = (n_1\, n_2 \cdots n_r)$ the vector of occupancy numbers for the energy cells. Of course, $n_1 + \cdots + n_r = n$. We shall call $W(\mathbf{n})$ the number of ways n molecules can find their way into the kr cells to produce the occupancy vector $\mathbf{n}$. There are at least three answers to the occupancy question, depending upon the assumptions one makes about the molecules under study.

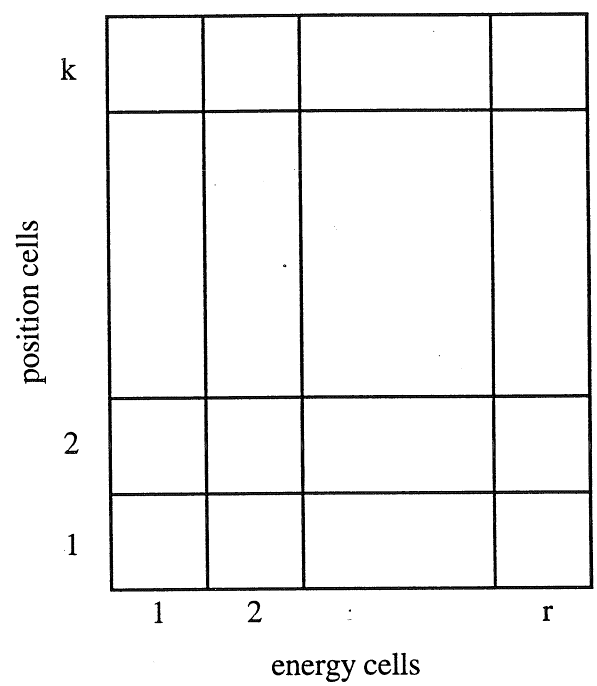

Figure 0.2. Phase space.

In our development of occupancy formulas, we shall break the distribution problem into two experiments. In the first, we distribute molecules over position cells, and in the second we distribute them over energy cells. We will count the number of ways of doing this as $W = W_1 W_2$, where W_1 is the number of ways to distribute over position and W_2 is the number of ways to distribute over energy.

Maxwell-Boltzmann. Maxwell and Boltzmann assume that the molecules are distinguishable and that an arbitrary number of them may occupy any position cell. The number of ways to distribute the n_i molecules in energy cell i into the k position cells is

$$W_{1i} = k^{n_i}. \tag{20}$$

The number of ways to distribute molecules from all energy cells into the position cells is

$$W_1 = \prod_{i=1}^{r} W_{1i} = \prod_{i=1}^{r} k^{n_i} = k^n. \tag{21}$$

The number of ways to distribute the original n distinguishable molecules into the r energy cells in the numbers $\mathbf{n}$ is

$$W_2(\mathbf{n}) = \binom{n}{n_1\, n_2 \cdots n_r}. \tag{22}$$

Therefore, the Maxwell-Boltzmann distribution is

$$W(\mathbf{n}) = W_1 W_2 = \binom{n}{n_1\, n_2 \cdots n_r} k^n. \tag{23}$$

This is just k^n times the multinomial distribution. Physicists know that this result is not useful for describing indistinguishable molecules, so they have applied an ad-hoc division by $n!$ to "make it work." This is called the "Boltzmann correction."

Bose-Einstein. Bose and Einstein assume that molecules are indistinguishable and that an arbitrary number may occupy any position cell. The number of ways to distribute the n_i indistinguishable molecules that reside in energy cell i into k position cells is

$$W_1 = \frac{k(k-1+n_i)!}{k_i! n_i!} = \binom{k-1+n_i}{n_i}. \tag{24}$$

This formula is really mysterious. Here is the way we explain it: we want to lay out a string of molecules and position cells with the understanding that the string will begin with a cell. There are k ways to choose the first cell. From then on, we want to permute the remaining $k-1$ cells and n_i molecules with the understanding that any molecules that immediately follow a cell occupy that cell. Two consecutive cells means that the first cell is unoccupied. This scheme is illustrated in Figure 0.3. This string of cells and molecules treats cells and molecules as distinguishable and ordered, but all $k!$ permutations of the cell-molecule pairs leaves the occupancy unchanged, as do all $n!$ permutations of the molecules within the cells. Thus, the occupancy number given in (24). The distribution over position cells is now

$$W_1 = \prod_{i=1}^{r} W_{1i} = \prod_{i=1}^{r} \binom{k-1+n_i}{n_i}. \tag{25}$$

The number of ways for indistinguishable molecules to occupy energy cells in the numbers $\mathbf{n}$ is

$$W_2(\mathbf{n}) = 1. \tag{26}$$

If this answer seems unclear, ask yourself how else you would achieve this occupancy when every molecule is indistinguishable from every other. The Bose-Einstein distribution of molecules over energy cells is

$$W(\mathbf{n}) = (1) \prod_{i=1}^{r} \binom{k-1+n_i}{n_i}. \tag{27}$$

$$\underline{c_{(1)}\ m_{(1)}\ c_{(2)}\ m_{(2)}\ m_{(3)}\ m_{(4)}\ c_{(3)}\ c_{(4)}\ m_{(5)}\ c_{(6)}\ c_{(7)}\ c_{(8)}\ m_{(6)}\ \cdots\ c_{(k)}}$$

Figure 0.3: *Laying out a sequence of position cells and molecules. (Cell (1) has molecule (1), cell (3) is empty, and so on.)*

Fermi-Dirac. Fermi and Dirac assume that molecules are indistinguishable and that a Pauli exclusion principle restricts occupancy of position cells to one or none. There are

$$W_{1i} = \binom{k}{n_i} \tag{28}$$

ways to distribute the n_i molecules in energy cell i into the k position cells. This is just the number of subsets of size n_i (the occupied cells) that can be constructed from a set of k cells. The distribution over position cells is now

$$W_1 = \prod_{i=1}^{r} W_{1i} = \prod_{i=1}^{r} \binom{k}{n_i}. \tag{29}$$

As in the case of Bose-Einstein, the number of ways to distribute indistinguishable molecules into energy cells with occupancy n is

$$W_2(\mathbf{n}) = 1. \tag{30}$$

The Fermi-Dirac distribution over energy cells is

$$W(\mathbf{n}) = (1)\prod_{i=1}^{r}\binom{k}{n_i}. \tag{31}$$

Problems:

In some of the problems of this section, the "probability" of some event is asked for. Use the following definition of probability for these problems, which is based on counting. In each case, the total number of possibilities is finite. Call this number N. Out of these there are M possible outcomes which satisfy the description of the event in question. The probability of the event is, by definition,

$$P = \frac{M}{N}.$$

0.1 There are ten pairs of shoes in a closet. You select four shoes. How many ways are there to get four shoes? How many ways are there to get exactly one pair? [Hint: This is a poker problem with ten faces (shoe types) and two suits (left and right).]

0.2 This article appeared in the July 24, 1996, issue of the *Fort Collins Coloradoan.* What is the probability of hitting 8-0-0 straight, and what should a fair payout be? What is the probability of hitting the box on 8-0-0, and what would a fair payout be?

> (By the Associated Press) HARTFORD,CT. Lottery players in Connecticut gambled on tragedy over the weekend and ended up breaking the bank. On Saturday, thousands of people played the daily lottery and put their money down on number 8-0-0, the flight number of the TWA jumbo jet that crashed into the Atlantic Ocean last week. When that number hit, lottery officials said they had to pay out just over \$1 million in winnings, about three times what was wagered. Of those who bought tickets for Saturday's drawing, 3,026 people bet the "800" number and another 3,182 chose the number in a three-way box, which means they would win with 800, 080, or 008. Gamblers who spent 50 cents on a straight bet won \$250, while people with three-way plays won \$83.50.

0.3 A student asked his teacher the following question:

> "You stated that the odds were the same whether or not the person playing the lotto stayed with the same numbers or picked different ones each week. This logic does not sit well with me. How can the odds of two random

> number generators picking the same six numbers be the same as one random number generator picking a set of six numbers?"

The teacher responded as follows:

> "You and I are going to match coins. The possible outcomes are you H me H, you H me T, you T me H, and you T me T. Two of these four possible outcomes produce matches. Now let's play the game this way: I show you my T, and you try to match it with a flip. One of your two outcomes produces a match. Same odds. In the lotto we generalize this to the flipping of a coin with 5,245,786 sides. If we are both flipping this 5,245,786-sided coin, there are this number squared of possible outcomes, 5,245,786 of which produce a match. The odds are 5,245,786 to 1. If, instead, I show you my number and you try to match it, you have one chance out of 5,245,786 to do it. Same odds."

What do you think of this answer? Let's assume you like it. Then why doesn't the State simply publish a good Lotto number and let people randomly draw six random numbers from a set of 42? Maybe the answer has more to do with human nature than with probability.

0.4 Fill out a table of odds and payouts for k matches in a fair $(42, 6)$ Lotto game. Then get a Lotto ticket and compare these fair payouts with what the State does.

0.5 Fill out a table of odds and payouts for k matches in a fair $(60, 20, 10)$ Keno game. Then get a Keno ticket and compare these payouts with what the State does.

0.6 A perfect bridge hand consists of all thirteen cards from one suit. How many perfect bridge hands are there? What are the odds of getting one?

0.7 For each of the poker hands described in Section 0.4, compute the number of ways to get it and the odds of getting it. Check that your numbers for hands (i) through (vi) add up to $\binom{52}{5}$.

0.8 The probability of getting exactly $n/2$ ones and $n/2$ zeros in an n-bit transmission is $\binom{n}{n/2}/2^n$ (namely, the ratio of the number of sequences that have this balance to the number of sequences). Compute this probability for illustrative values of n. Use Sterling's approximation,

$$n! \cong \sqrt{2\pi n}\, n^n e^{-n},$$

to find the limiting probability. Evidently we do not tend to see balanced sequences, so what do we mean when we say that sequences are roughly balanced? Make your answer quantitative.

0.9 Prove that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

0.10 The binomial coefficient $\binom{n}{k}$ is well defined if $n \geq k > 0$. Extend this to

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!}, & 0 \leq k \leq n \\ 0, & k > n \end{cases}$$

where we set $0! = 1$. The binomial coefficient is left undefined for n and k not integers or negative numbers. Let $\mathbf{C}$ and $\mathbf{A}$ be $(n+1) \times (n+1)$ matrices whose elements are given by

$$C_{ij} = \binom{i+j}{i} \quad \text{for} \quad 0 \leq i, j \leq n$$

$$A_{ij} = \binom{i}{j} \quad \text{for} \quad 0 \leq i, j \leq n.$$

Prove that $\mathbf{C} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$. As an example,

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{bmatrix}; \qquad \mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}.$$

(One can see the Pascal triangle in each matrix.)

0.11 Let's say that 2,000 voters in a small town are equally divided between men and women, with 56% of the women Democrats and 52% of the men Republicans. What is the probability that a random draw of 200 voters will produce these percentages?

0.12 Natalie's father wants to encourage her interest in tennis. He offers her a trip to Wimbledon if she wins at least two consecutive matches in succession against himself (a mediocre player) and the club champion (a fine player) in a three-match tournament of alternate matches (either f-cc-f or cc-f-cc). Should she play her father twice or the club champion twice?

0.13 Seven major league teams (Rockies, Dodgers, Giants, Cardinals, Reds, Cubs, and Mets) divide up 63 players into nine-person teams. How many ways are there to do this?

0.14 How many ways are there to divide up the 63 players into nine-person teams if team name does not matter? (Hint: if nine are on the same team, it does not matter what the team name is.)

0.15 In a class of 63 students, there are 50 men and 13 women. What is the probability of getting 7 women in a random draw of 32 students?

0.16 I am running a fair Keno game. There are ten numbers at play, four of which are winners and six of which are losers. For $1, you buy three numbers. How much should you win if you match two of the good numbers?

0.17 An encounter group has six couples. What is the probability of selecting two couples in a random selection of five people?

0.18 Are you more likely to get four ones in eight binary transmissions or five ones in nine transmissions? Explain.

0.19 There are eight teams of three players each registered for a 3-on-3 basketball tournament. I am going to randomly draw six players. What is the probability that at least one of the registered teams shows up in my random selection of six players? (Hint: consider xxxyzw, xxxyyz, and xxxyyy.)

0.20 For manageable values of $k > n$, $n > r$, and r, determine the Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac answers for the number of ways to *equally* distribute n molecules over r energy cells.