Chapter 3. Combinatorics and examples.

3.1 Fundamental rule of counting.

Theorem (Fundamental rule of counting). *Given m objects $a_1, \ldots, a_m$ and n objects $b_1, \ldots, b_n$, there are mn ordered pairs of the form (a_i, b_j). Alternate statement: If an experiment consists of two tasks the first of which can be performed in m ways and, regardless of the particular outcome of the first task, a second task which can be performed in n ways, then the experiment itself can be performed in mn ways.*

Proof. Let $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ be given. Clearly there are n ordered pairs of the form (a_1, b_j). Thus, by letting the first element vary through its m possible values there are mn ordered pairs of the form (a_i, b_j). □

Corollary. *There are $n_1 n_2 \cdots n_k$ ways in which an ordered k–tuple can be formed when there are n_1 choices for the first element, n_2 choices for the second element, ..., and n_k choices for the k^{th} element.*

Example. Tossing a pair of dice. Suppose that a pair of fair dice is tossed and, for $x = 2, 3, \ldots, 12$, let p_x denote the probability that the sum of the numbers on the upturned faces is x. The elementary outcomes for this experiment can be represented by the 36 ordered pairs of the form (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$. Since the dice are assumed to be fair we will assume that these 36 elementary outcomes are equally probable. The sums corresponding to these elementary outcomes are indicated in the following table.

Sums.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Hence, $p_2 = p_{12} = 1/36$, $p_3 = p_{11} = 2/36$, $p_4 = p_{10} = 3/36$, $p_5 = p_9 = 4/36$, $p_6 = p_{11} = 5/36$, and $p_7 = 6/36$.

3.2. Ordered samples.

Consider a population $\{a_1, \ldots, a_N\}$ containing N distinct objects. An ordered sample of size n is an ordered collection of n objects selected from the N objects in the population. It is helpful to envision selecting the n objects which comprise the sample one at a time.

There are two ways that such a sample can be selected. The elements of the sample can be selected with replacement, so that at each step the element is selected from the entire population, or the elements of the sample can be selected without replacement, so that at each step an object is removed from the population and there is one less object in the population for the next selection. When sampling with replacement a particular object can appear more than once in the sample and there is no restriction on the size n of the sample. When sampling without replacement no particular object can appear more than once in the sample and the sample size n clearly cannot exceed the population size N.

Given a population of size N consider the selection of an ordered sample of size n as a sequence of n steps where the i^{th} element of the sample is selected at step i. When sampling with replacement there are N choices at each of the n steps. Thus there are N^n possible ordered samples of size n when the sample is selected with replacement from a population of size N. When sampling without replacement there are N choices at the first step, $N-1$ choices at the second step, and so on, with $N-n+1$ choices at the n^{th} step (assuming that $n \leq N$). Thus, assuming that $n \leq N$, there are $N(N-1)\cdots(N-n+1)$ possible ordered samples of size n when the sample is selected without replacement from a population of size N. Note that each of these counts, N^n and $N(N-1)\cdots(N-n+1)$, is a product of n integers. For ease of reference these results are summarized in the following theorem.

Theorem (ordered samples). *Let a population of size N be given. There are N^n possible ordered samples of size n when the sample is selected with replacement. For $n \leq N$, there are $N(N-1)\cdots(N-n+1)$ possible ordered samples of size n when the sample is selected without replacement.*

The process of selecting an ordered sample of size n with replacement from a population of N distinct objects is equivalent to the process of placing n balls into N boxes when placement of more than one ball into the same box is allowed. This equivalence is clear if we identify each of the n steps involved in selecting the objects from the population with the assignment of a box to a ball. Similarly, the process of selecting an ordered sample of size n without replacement from a population of N distinct objects is equivalent to the process of placing n balls into N boxes when placement of more than one ball into the same box is not allowed.

Note that if the sample is selected without replacement and $n = N$, then the ordered sample is a permutation (ordered arrangement) of the objects $\{a_1, \ldots, a_N\}$ which comprise the population. There are $N! = N(N-1)\cdots 1$ (read this as N factorial) permutations of N objects. By convention $0! = 1$. With this convention, for $1 \leq n \leq N$, the number of ordered samples of size n selected without replacement from a population of N objects (the number of permutations of N objects taken n at a time) is $\frac{N!}{(N-n)!}$.

3.3. Subpopulations and partitions.

As above a population of size N is an unordered collection (set) $\{a_1, \ldots, a_N\}$ of N distinct objects. Two populations are said to be different when one contains an object which does not belong to the other. Given a population $\{a_1, \ldots, a_N\}$, a subpopulation of size n is an unordered collection of n distinct a_i values, *i.e.*, a subset of size n. We will now determine, for fixed $n < N$, the number of subpopulations of size n that can be formed from the objects in a population of size N. We already know that there are $N(N-1)\cdots(N-n+1)$ ordered samples of size n selected without replacement from a population of size N. Since each of these ordered samples of size n can be ordered in $n!$ ways, *i.e.*, there are $n!$ permutations of a particular set of n distinct objects, it follows that the number of subpopulations of size n (selected from a population of size N) is given by the binomial coefficient

$$\binom{N}{n} = \frac{N(N-1)\cdots(N-n+1)}{n(n-1)\cdots 1} = \frac{N!}{n!(N-n)!}.$$

Note that in the ratio of products expression for $\binom{N}{n}$ the numerator and denominator products are both products of n integers. For ease of reference this result is summarized in the following theorem.

Theorem (subpopulations). *Let a population of size N be given. For $n \le N$, the number of subpopulations of size n is given by the binomial coefficient*

$$\binom{N}{n} = \frac{N(N-1)\cdots(N-n+1)}{n(n-1)\cdots 1} = \frac{N!}{n!(N-n)!}.$$

A subpopulation of size n selected from a population of size N is also said to be a combination of N objects taken n at a time. Thus the binomial coefficient $\binom{N}{n}$ is the number of combinations of N objects taken n at a time. Note also that selecting a subpopulation of size n from a population of size N is equivalent to selecting a collection of $N-n$ objects which are excluded from the subpopulation. Thus, recalling the convention $0! = 1$, which leads to the convention $\binom{N}{0} = \binom{N}{N} = 1$, for $0 \le n \le N$, we have

$$\binom{N}{n} = \binom{N}{N-n}.$$

As noted above the selection of a subpopulation of size n from a population of size N is equivalent to partitioning (dividing) the population into two subpopulations, one of size n and one of size $N-n$. For $m \ge 2$ and $n_1, \ldots, n_m$ such that $n_1 + \cdots + n_m = N$, we might ask: In how many ways can a population of size N be partitioned into m subpopulations of respective sizes $n_1, \ldots, n_m$? To answer this question consider the formation of such a partition via a sequence of m steps. At the first step there are N objects to choose from and $\binom{N}{n_1}$ ways to select the first subpopulation. At the second step there are $N-n_1$ objects to choose from and $\binom{N-n_1}{n_2}$ ways to select the second subpopulation. Continuing with this argument when we reach the last (m^{th}) step we find that there are $N-n_1-\cdots-n_{m-1}=n_m$ objects to choose from and $\binom{n_m}{n_m}=1$ way to select the final subpopulation. The final answer to our question is given in the following theorem.

Theorem (partitions). *Let a population of size N be given. For $m \geq 2$ and $n_1, \ldots, n_m$ such that each $n_i \geq 1$ and $n_1+\cdots+n_m=N$, the number of partitions of the population into m subpopulations of respective sizes $n_1, \ldots, n_m$ is given by the multinomial coefficient*

$$\binom{N}{n_1}\binom{N-n_1}{n_2}\binom{N-n_1-n_2}{n_3}\cdots\binom{n_m}{n_m}=\frac{N!}{n_1!n_2!\cdots n_m!}.$$

Note that if we allow $n_i=0$ in this theorem, then the expression is still valid but the number of nontrivial subpopulations in the partition is reduced by the number of i for which $n_i=0$.

Example. Binomial distribution. Suppose that n balls are placed at random and independently into N boxes. Here the word independently indicates that more than one ball may be placed into the same box and the word random indicates that we are going to assume that the N^n elementary outcomes, represented as n–tuples with elements taking values in $\{1,\ldots,N\}$, are equally likely. For an integer x, with $0 \leq x \leq n$, we will find the probability p_x that exactly x balls are placed in box 1. First note that an outcome is favorable for this event when it contains exactly x ones and exactly $n-x$ values not equal to one, since these are the elementary outcomes with x balls in box 1 and $n-x$ balls in other boxes. We know that there are $\binom{n}{x}$ partitions of a population of n objects (the balls) into 2 subpopulations of respective sizes x and $n-x$. For a specific partition of this form the n–tuple must have ones in the specified x positions and values not equal to one in the other positions. Since there is only one way to choose a one and there are $N-1$ ways to choose an value other than one, there are $1^x(N-1)^{n-x}=(N-1)^{n-x}$ elementary outcomes with ones in the x positions for the specific partition and values other than one in the others. Hence, there are $\binom{n}{x}(N-1)^{n-x}$ elementary outcomes that are favorable for the event "exactly x balls are placed in box 1" and, for $x=0,1,\ldots,n$, the probability that exactly x balls of the n balls are placed in box 1 is

$$p_x=\binom{n}{x}\frac{(N-1)^{n-x}}{N^n}=\binom{n}{x}\left(\frac{1}{N}\right)^x\left(\frac{N-1}{N}\right)^{n-x}.$$

Note that, if we consider the random placement of a single ball into one of N boxes, then $\frac{1}{N}$ is the probability that the ball is placed in box 1 and $\frac{N-1}{N}$ is the probability that it is placed in a box other than box 1.

We can modify this example to get a more general expression for p_x that involves probabilities of the form $\frac{N_1}{N}$ and $\frac{N-N_1}{N}$. To do this simply assume that boxes 1 through N_1 are one color (say red) and boxes N_1+1 through N are another color (say green); redefine the event as "exactly x balls are placed in a red box"; and, modify the argument to count the number of ways x balls can be placed in a red box and the others in a green box, yielding $N_1^x(N-N_1)^{n-x}$ instead of $1^x(N-1)^{n-x}$. With this modification, for $x=0,1,\ldots,n$, the probability that exactly x balls of the n balls are placed in a red box is

$$p_x = \binom{n}{x}\frac{N_1^x(N-N_1)^{n-x}}{N^n} = \binom{n}{x}\left(\frac{N_1}{N}\right)^x\left(\frac{N-N_1}{N}\right)^{n-x}.$$

In this case, if we consider the random placement of a single ball into one of N boxes, then $\frac{N_1}{N}$ is the probability that the ball is placed in a red box and $\frac{N-N_1}{N}$ is the probability that it is placed in a green box.

As noted above, the process of placing n balls into N boxes, when placement of more than one ball into the same box is allowed, is equivalent to the process of selecting an ordered sample of size n with replacement from a population of N distinct objects. Thus this expression for p_x can also be viewed as the probability of obtaining a sample which contains exactly x objects of one type (red) when a random sample of size n is selected with replacement from a population of N objects of which N_1 are of one type (red) and $N-N_1$ are of a second type (green). In this context the adjective random indicates that each of the possible samples is equally probable.

Example. If a fair die is tossed 10 times, then the probability of observing exactly x ones is $p_x = \binom{10}{x}(\frac{1}{6})^x(\frac{5}{6})^{10-x}$. For example, the probability of observing exactly 3 ones is $\binom{10}{3}(\frac{1}{6})^3(\frac{5}{6})^7 = 120 \cdot 5^7/6^{10} \approx .1550$.

For completeness we should verify that the binomial probabilities $p_0, p_1, \ldots, p_n$ sum to one. This follows from the binomial theorem with $a=p$ and $b=1-p$, where p denotes the probability of obtaining a "red" ball when a single ball is selected at random.

Binomial theorem. *Given real numbers a and b and a positive integer n*

$$(a+b)^n = \sum_{x=0}^{n}\binom{n}{x}a^x b^{n-x}.$$

Proof. Let the real numbers a and b and the positive integer n be given. Multiplying the n terms in the product $(a+b)^n$ yields a sum of 2^n products of the form $a^x b^{n-x}$ for values of x between 0 and n. For each value of x, there are $\binom{n}{x}$ ways in which the expression $a^x b^{n-x}$ appears, since these expression correspond to the various combinations obtained by choosing x $(a+b)$ terms to contribute an a (and $n-x$ $(a+b)$ terms to contribute a b). Thus appropriate grouping of these products yields the result. □

Example. Hypergeometric distribution. Suppose that a sample of n balls (objects) is selected at random without replacement from a population of N balls of which N_1 are red (of one type) and $N_2 = N - N_1$ are green (of a second type). In this context the word random indicates that we assume that the $\binom{N}{n}$ samples of size n, represented as (unordered) sets of size n, are equally probable. For an integer x, with $0 \le x \le n$, we will find the probability p_x that the sample contains exactly x red balls (and consequently exactly $n-x$ green balls). First note that, since the population contains N_1 red balls and $N_2 = N - N_1$ green balls and we are sampling without replacement, we must have $p_x = 0$ when $x > N_1$ and when $n - x > N - N_1$. For admissible values of x there are $\binom{N_1}{x}$ ways to select x red balls for the sample and $\binom{N-N_1}{n-x}$ ways to select $n-x$ green balls for the sample. Thus there are $\binom{N_1}{x}\binom{N-N_1}{n-x}$ outcomes favorable for the event "the sample contains exactly x red balls" and, for admissible values of x, the probability of this event is

$$p_x = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}}.$$

In the context of placing of balls into boxes, this p_x is the probability that exactly x balls are placed in a red box when n balls are placed at random into N boxes, of which N_1 are red and $N - N_1$ are green, subject to the restriction that no more than one ball can be placed in a box.

Example. If a poker hand is dealt randomly, then the probability that the hand contains exactly x aces is $p_x = [\binom{4}{x}\binom{48}{5-x}]/\binom{52}{5}$. For example, the probability that the hand contains exactly two aces is $[\binom{4}{2}\binom{48}{3}]/\binom{52}{5} = 6 \cdot 17296/2598960 \approx .0399$

For completeness we should should verify that the hypergeometric probabilities p_0, $p_1, \ldots, p_n$ sum to one. To simplify the argument we will adopt the convention that $\binom{a}{b} = 0$ whenever $b > a$. Consider the number of combinations of N balls taken n at a time, when N_1 are red and $N - N_1$ are green. We know that there there are $\binom{N}{n}$ combinations altogether. For each value of w between 0 and n, there are $\binom{N_1}{w}\binom{N-N_1}{n-w}$ combinations with w red balls and $n-w$ green balls (recall that for some values of w there may not be any such combinations and this product is zero). Since one of the combinations of w red and $n-w$ green balls must occur we get $\sum_{w=0}^{n}\binom{N_1}{w}\binom{N-N_1}{n-w} = \binom{N_1}{0}\binom{N-N_1}{n} + \binom{N_1}{1}\binom{N-N_1}{n-1} + \cdots + \binom{N_1}{n}\binom{N-N_1}{0} = \binom{N}{n}$ which shows that $p_0, p_1, \ldots, p_n$ sum to one.

We will now extend these arguments to find similar probabilities corresponding to random samples from a population comprised of object of three or more types.

Example. Multinomial distribution. The binomial distribution is readily generalized to allow for balls of three or more types. Suppose that a sample of n balls is selected at random with replacement from a population of $N = N_1 + N_2 + N_3$ balls of which N_1 are of type one, N_2 are of type two, and N_3 are of type three. For $n = n_1 + n_2 + n_3$, we will find the probability p_{n_1,n_2,n_3} that the sample contains n_1 balls of type one, n_2 balls of type two, and n_3 balls of type three. The N^n elementary outcomes, represented by n–tuples in which a value can occur more than once, are equally probable. There are $\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}$ ways in which we can choose positions in an n–tuple for n_1 balls of type one, n_2 balls of type two, and n_3 balls of type three. For each particular choice of positions there are $N_1N_2N_3$ such elementary outcomes. Thus, for $n = n_1 + n_2 + n_3$, the probability of the event "the sample contains n_1 balls of type one, n_2 balls of type two, and n_3 balls of type three" when the random sample of n balls is selected with replacement and $N = N_1 + N_2 + N_3$ is

$$p_{n_1,n_2,n_3} = \binom{n}{n_1}\binom{n-n_1}{n_2}\left(\frac{N_1}{N}\right)^{n_1}\left(\frac{N_2}{N}\right)^{n_2}\left(\frac{N_3}{N}\right)^{n_3},$$

since $\binom{n-n_1-n_2}{n_3} = \binom{n_3}{n_3} = 1$. The extension to balls of four of more types should be obvious.

Example. If a fair die is tossed 10 times, then the probability of observing exactly 3 ones and exactly 4 twos is $\binom{10}{3}\binom{7}{4}\binom{3}{3}(\frac{1}{6})^3(\frac{1}{6})^4(\frac{4}{6})^4 = [120 \cdot 35 \cdot 1 \cdot 4^4]/6^{10} \approx .0178$.

Example. Multiple hypergeometric distribution. The hypergeometric distribution is also readily generalized to allow for balls of three or more types. Suppose that a sample of n balls is selected at random without replacement from a population of $N = N_1 + N_2 + N_3$ balls of which N_1 are of type one, N_2 are of type two, and N_3 are of type three. For $n = n_1 + n_2 + n_3$, we will find the probability p_{n_1,n_2,n_3} that the sample contains n_1 balls of type one, n_2 balls of type two, and n_3 balls of type three. The $\binom{N}{n}$ elementary outcomes, represented by (unordered) sets of size n, are equally probable. In this case we must have $n_1 \le N_1$, $n_2 \le N_2$, and $n_3 \le N_3$. Assuming that this is true there are $\binom{N_1}{n_1}\binom{N_2}{n_2}\binom{N_3}{n_3}$ elementary outcomes that are favorable for the event of interest. Thus, for admissible values of n_1, n_2, and n_3, the probability of the event "the sample contains n_1 balls of type one, n_2 balls of type two, and n_3 balls of type three" when the random sample of n balls is selected without replacement and $N = N_1 + N_2 + N_3$ is

$$p_{n_1,n_2,n_3} = \frac{\binom{N_1}{n_1}\binom{N_2}{n_2}\binom{N_3}{n_3}}{\binom{N}{n}}.$$

Again the extension to balls of four of more types should be obvious.

Example. If a poker hand is dealt randomly, then the probability that the hand contains exactly 2 aces and exactly 2 kings is $[\binom{4}{2}\binom{4}{2}\binom{44}{1}]/\binom{52}{5} \approx .0006$.

Example. Geometric distribution. Suppose balls (objects) are selected at random with replacement from a population of N balls, of which N_1 are red (of one type) and $N_2 = N - N_1$ are green (of a second type), sequentially until a red ball is selected and let n denote the number of the selection on which the red ball occurs. In this context the elementary outcomes can be represented by sequences of the form R, GR, GGR, ..., with R representing a red ball and G a green ball. In this example the sample space is countably infinite. We will find the probability p_n that the first red ball occurs on the n^{th} selection. For a fixed value of n the elementary outcomes which are favorable for the event "the first red ball occurs on the n^{th} selection" are of the form $G \dots GR$ with $n-1$ $G's$. Assume that the N balls are distinguishable (if not they can always be suitably numbered), then there are $N_2^{n-1}N_1$ elementary outcomes which are favorable for the event of interest and there are N^n possible elementary outcomes. Thus, for $n = 1, 2, \dots$,

$$p_n = \frac{N_2^{n-1}N_1}{N} = \left(\frac{N_2}{N}\right)^{n-1}\left(\frac{N_1}{N}\right) = q_1^{n-1}p_1,$$

where $p_1 = \frac{N_1}{N}$ and $q_1 = 1 - p_1$. To see that this is a valid assignment of probabilities we need to verify that $\sum_{n=1}^{\infty} p_n = 1$. Note that $\sum_{n=1}^{\infty} p_n = p_1 \sum_{n=0}^{\infty} q_1^n = p_1/(1-q_1) = 1$, since the geometric series $\sum_{n=0}^{\infty} q_1^n$ converges to $1/(1-q_1)$. These p_i are often labeled in terms of the number of green balls selected before the first red ball is selected, *i.e.*, in terms of $n-1$ instead of n.

Example. If a fair die is tossed repeatedly until a one appears, then the probability of observing this initial one on the fifth toss is $(\frac{5}{6})^4(\frac{1}{6}) \approx .0804$.

Example. Matching. Suppose that two decks of N distinguishable cards are placed in random order and compared. If the two cards in a particular location are the same, then we have a match. You can also think of placing N numbered balls into N numbered boxes and looking for matches. We will find the probability P_1 that there is at least one match. This scenario can be rephrased in many amusing ways. For example: What is the probability that at least one letter will be placed in the correct envelope if N letters are placed at random into N envelopes? Letting A_i denote the event that there is a match at position i ($i = 1, \dots, N$) note that $P_1 = \Pr(A_1 \cup \cdots \cup A_N)$. For counting purposes we can imagine labeling the cards, as $1, \dots, N$, in both decks to agree with the way that they are currently ordered in the first deck and we can represent an elementary outcome as a permutation of $1, \dots, N$ indicating their ordering in the second deck. The assumption of random ordering is assumed to mean that each of these permutations has probability $\frac{1}{N!}$. For a specified position i ($i = 1, \ldots, N$) there are $(N-1)!$ permutations with a match at position i. Thus $\Pr(A_i) = \frac{(N-1)!}{N!}$. Similarly, for two distinct specified positions i and j, there are $(N-2)!$ permutations with matches at positions i and j. Thus $\Pr(A_i A_j) = \frac{(N-2)!}{N!}$. In general, for k distinct specified positions, there are $(N-k)!$ permutations with matches at these k positions. Thus, letting $i_1, \ldots, i_k$ indicate the k positions, $\Pr(A_{i_1} A_{i_2} \cdots A_{i_k}) = \frac{(N-k)!}{N!}$. We will now apply Theorem 2.10 to find P_1. Letting $S_1 = \sum_i \Pr(A_i)$, $S_2 = \sum_{i<j} \Pr(A_i A_j)$, and so on Theorem 2.10 indicates that $P_1 = S_1 - S_2 + S_3 - S_4 + \cdots \pm S_N$. Since there are $\binom{N}{k}$ ways to select k distinct positions we see that $S_k = \binom{N}{k} \frac{(N-k)!}{N!} = \frac{1}{k!}$. Hence,

$$P_1 = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots \pm \frac{1}{N!}.$$

It is interesting to consider how this probability depends on the value of N. For example, how does the probability of at least one match with a deck of 7 cards compare to the probability for a deck of 7,000 cards? We shall see that it turns out that this probability is nearly independent of the value of N and roughly equal to $\frac{2}{3}$. The values of P_1 for $N = 3$ through 7 are:

$N =$	3	4	5	6	7
$P_1 =$	$\frac{2}{3} \approx .66667$	$\frac{15}{24} \approx .62500$	$\frac{76}{120} \approx .63333$	$\frac{455}{720} \approx .63194$	$\frac{3186}{5040} \approx .63214$

For larger values of N, note that $-P_1$ corresponds to the first N terms of the expansion

$$e^{-1} = -1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \cdots$$

Hence for large N we have the approximation $P_1 \approx 1 - e^{-1} \approx .63212$. Referring to the table above we see that this approximation is quite accurate for relatively small N with agrement to 4 decimal places when $N = 7$.