1.4 How to count

Unit 1: Probability and Counting

Adapted from Blitzstein-Hwang Chapter 1.

Calculating the naive probability of an event A involves counting the number of pebbles in A and the number of pebbles in the sample space S. Often the sets we need to count are extremely large, so it would be tedious or infeasible to count the possibilities one by one. To address this challenge, we will introduce a few techniques for counting.

Multiplication rule

In some problems, we can directly count the number of possibilities using a basic but versatile principle called the *multiplication rule*. We'll see that the multiplication rule leads naturally to counting rules for *sampling with replacement* and *sampling without replacement*, two scenarios that often arise in probability and statistics.

> **Theorem 1.4.1 (Multiplication rule).**
>
> Consider a compound experiment consisting of two sub-experiments, Experiment A and Experiment B. Suppose that Experiment A has a possible outcomes, and for each of those outcomes Experiment B has b possible outcomes. Then the compound experiment has ab possible outcomes.

To see why the multiplication rule is true, imagine a tree diagram as in Figure 1.4.2. Let the tree branch a ways according to the possibilities for Experiment A, and for each of those branches create b further branches for Experiment B. Overall, there are ab possibilities.

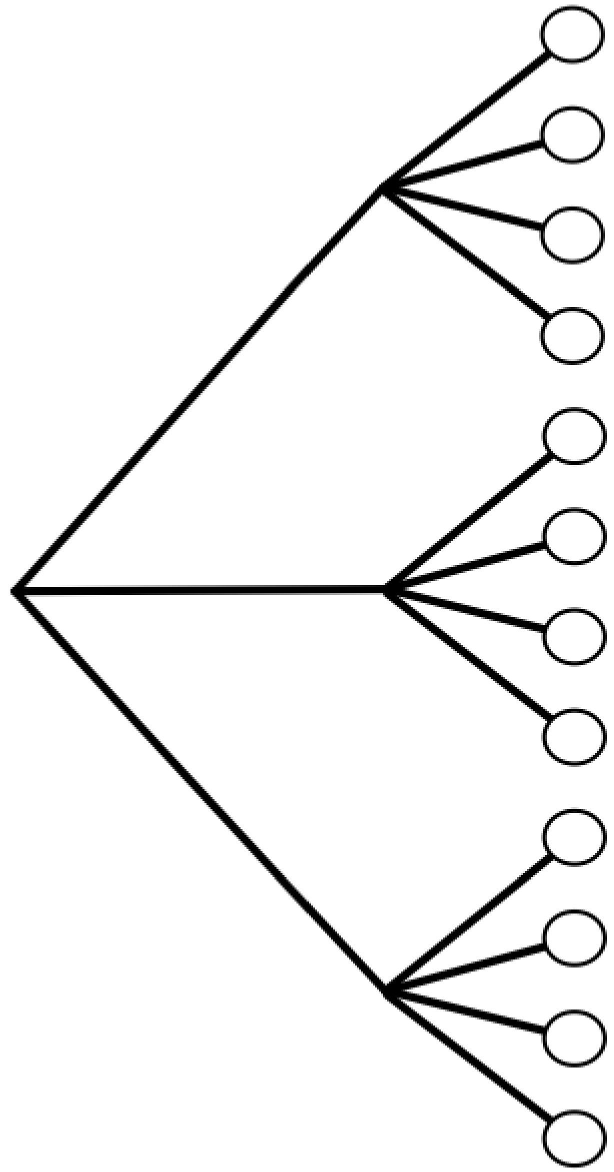

Figure 1.4.2: Tree diagram illustrating the multiplication rule. If Experiment A has 3 possible outcomes and Experiment B has 4 possible outcomes, then overall there are $3 \cdot 4 = 12$ possible outcomes.

View Larger Image

Image Description

Example 1.4.3 (Ice cream cones).

Suppose you are buying an ice cream cone. You can choose whether to have a cake cone or a waffle cone, and whether to have chocolate, vanilla, or strawberry as your flavor. This decision process can be visualized with a tree diagram, as in Figure 1.4.4.

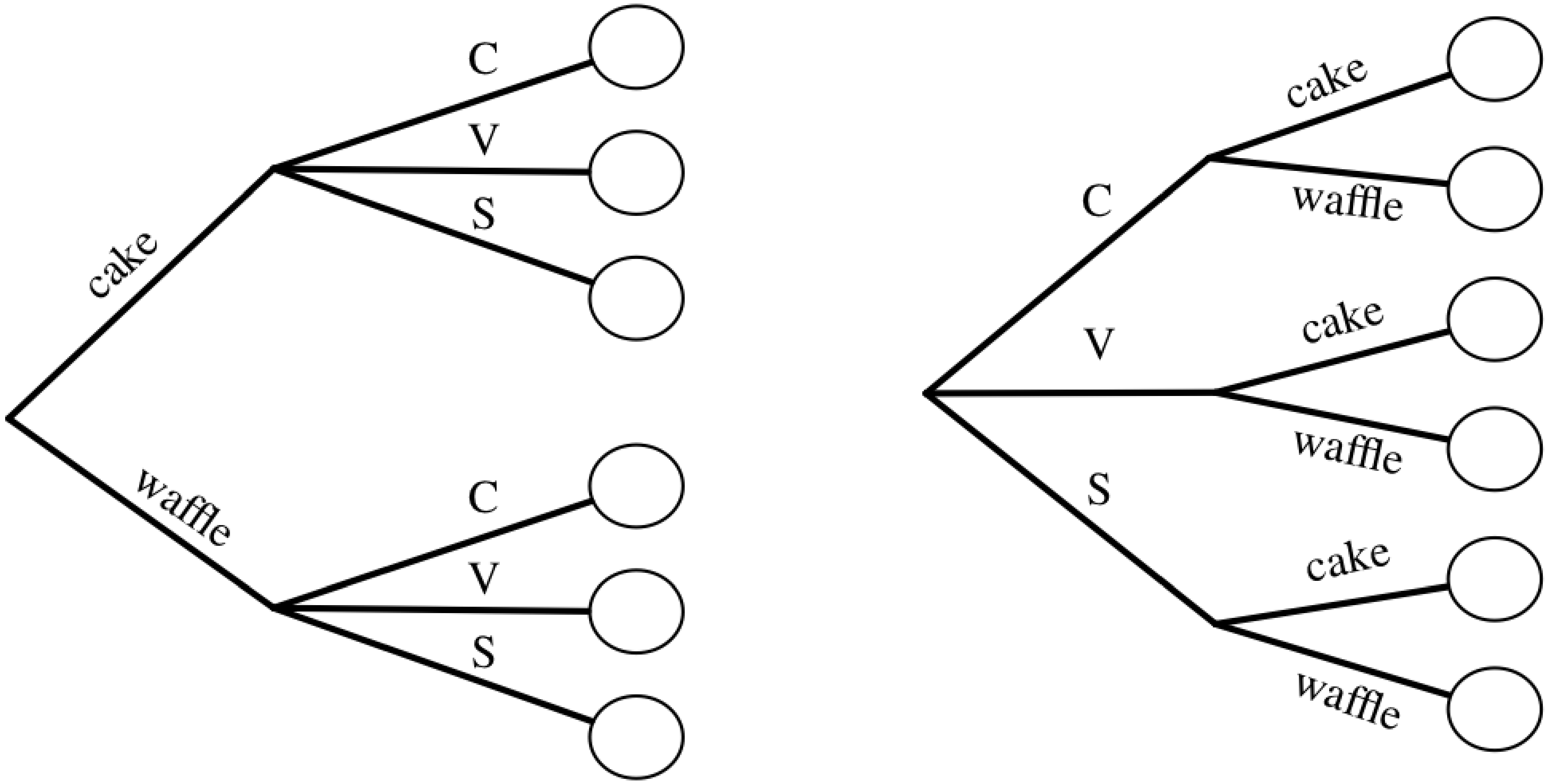

Figure 1.4.4: Tree diagram for choosing an ice cream cone. Regardless of whether the type of cone or the flavor is chosen first, there are $2 \cdot 3 = 3 \cdot 2 = 6$ possibilities.

View Larger Image
Image Description

By the multiplication rule, there are $2 \cdot 3 = 6$ possibilities. This is a very simple example, but is worth thinking through in detail as a foundation for thinking about and visualizing more complicated examples. Soon we will encounter examples where drawing the tree in a legible size would take up more space than exists in the known universe, yet where conceptually we can still think in terms of the ice cream example. Note that:

1. It doesn't matter whether you choose the type of cone first ("I'd like a waffle cone with chocolate ice cream") or the flavor first ("I'd like chocolate ice cream on a waffle cone"). Either way, there are $2 \cdot 3 = 3 \cdot 2 = 6$ possibilities.

2. It doesn't matter whether the same flavors are available on a cake cone as on a waffle cone. What matters is that there are exactly 3 flavor choices for each cone choice. If for some strange reason it were forbidden to have chocolate ice cream on a waffle cone, with no substitute flavor available (aside from vanilla and strawberry), there would be $3 + 2 = 5$ possibilities and the multiplication rule wouldn't apply. In larger examples, such complications could make counting the number of possibilities vastly more difficult.

We can use the multiplication rule to arrive at formulas for sampling with and without replacement. Many experiments in probability and statistics can be interpreted in one of these two contexts, so it is appealing that both formulas follow directly from the same basic counting principle.

THEOREM 1.4.5 (SAMPLING WITH REPLACEMENT).

Consider n objects and making k choices from them, one at a time *with replacement* (i.e., choosing a certain object does not preclude it from being chosen again). Then there are n^k possible outcomes.

For example, imagine a jar with n balls, labeled from 1 to n. We sample balls one at a time with replacement, meaning that each time a ball is chosen, it is returned to the jar. Each sampled ball is a sub-experiment with n possible outcomes, and there are k sub-experiments. So by the multiplication rule there are n^k ways to obtain a sample of size k.

THEOREM 1.4.6 (SAMPLING WITHOUT REPLACEMENT).

Consider n objects and making k choices from them, one at a time *without replacement* (i.e., choosing a certain object precludes it from being chosen again). Then there are $n(n-1)\cdots(n-k+1)$ possible outcomes, for $k \leq n$ (and 0 possibilities for $k > n$).

This result also follows directly from the multiplication rule: each sampled ball is again a sub-experiment, and the number of possible outcomes decreases by 1 each time. Note that for sampling k out of n objects without replacement, we need $k \leq n$, whereas in sampling with replacement the objects are inexhaustible.

Example 1.4.7 (Permutations and factorials).

A *permutation* of $1, 2, \dots, n$ is an arrangement of them in some order, e.g., $3, 5, 1, 2, 4$ is a permutation of $1, 2, 3, 4, 5$. By Theorem 1.4.6 with $k = n$, there are $n!$ permutations of $1, 2, \dots, n$. For example, there are $n!$ ways in which n people can line up for ice cream. (Recall that $n! = n(n-1)(n-2)\cdots 1$ for any positive integer n, and $0! = 1$.)

Theorems 1.4.5 and 1.4.6 are theorems about *counting*, but when the naive definition applies, we can use them to calculate *probabilities*. This brings us to our next example, a famous problem in probability called the *birthday problem*. The solution incorporates both sampling with replacement and sampling without replacement.

Example 1.4.8 (Birthday problem).

There are k people in a room. Assume each person's birthday is equally likely to be any of the 365 days of the year (we exclude February 29), and that people's birthdays are independent (we assume there are no twins in the room). What is the probability that two or more people in the group have the same birthday?

Solution

There are 365^k ways to assign birthdays to the people in the room, since we can imagine the 365 days of the year being sampled k times, with replacement. By assumption, all of these possibilities are equally likely, so the naive definition of probability applies.

Used directly, the naive definition says we just need to count the number of ways to assign birthdays to k people such that there are two or more people who share a birthday. But this counting problem is hard, since it could be Emma and Steve who share a birthday, or Steve and Naomi, or all three of them, or the three of them could share a birthday while two others in the group share a different birthday, or various other possibilities.

Instead, let's count the complement: the number of ways to assign birthdays to k people such that no two people share a birthday. This amounts to sampling the 365 days of the year *without* replacement, so the number of possibilities is $365 \cdot 364 \cdot 363 \cdots (365 - k + 1)$ for $k \leq 365$. Therefore the probability of no birthday matches in a group of k people is

$$P(\text{no birthday match}) = \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k},$$

and the probability of at least one birthday match is

$$P(\text{at least 1 birthday match}) = 1 - \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k}.$$

Figure 1.4.9 plots the probability of at least one birthday match as a function of k. The first value of k for which the probability of a match exceeds 0.5 is $k = 23$. Thus, in a group of 23 people, there is a better than 50% chance that two or more of them will have the same birthday. By the time we reach $k = 57$, the probability of a match exceeds 99%.

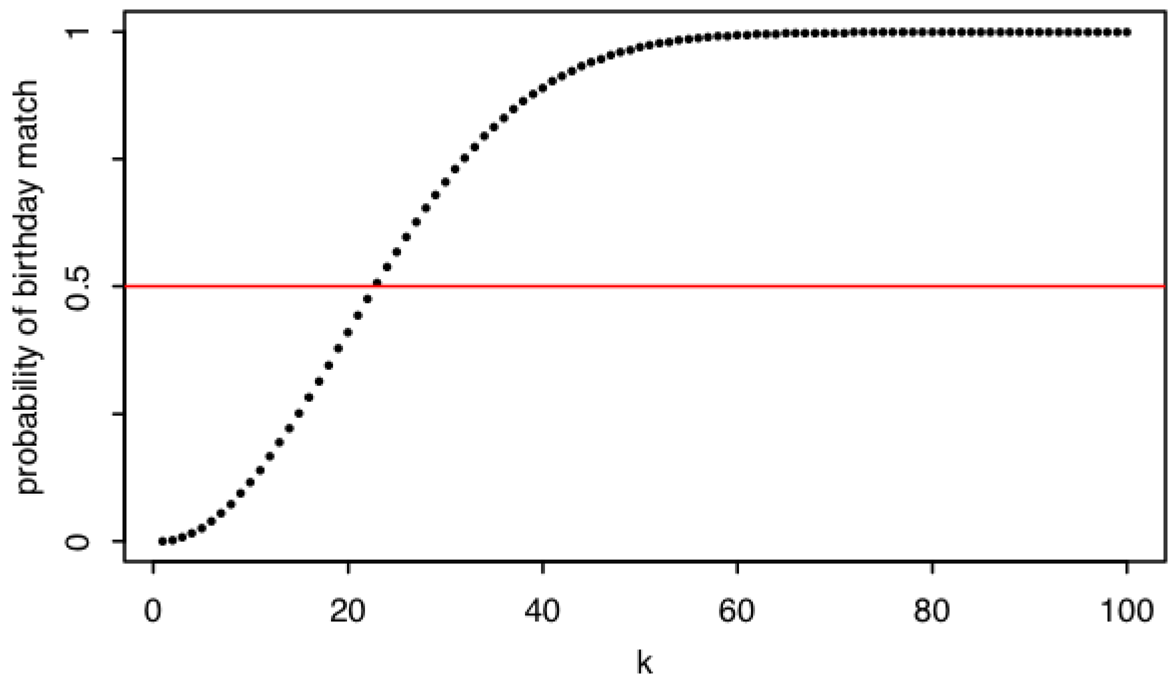

Figure 1.4.9: Probability that in a room of k people, at least two were born on the same day. This probability first exceeds 0.5 when $k = 23$.

View Larger Image

Image Description

Of course, for $k = 366$ we are *guaranteed* to have a match, but it's surprising that even with a much smaller number of people it's overwhelmingly likely that there is a birthday match. For a quick intuition into why it should not be so surprising, note that with 23 people there are $\binom{23}{2} = 253$ *pairs* of people, any of which could be a birthday match.

> **Warning 1.4.10 (Labeling objects).**
>
> Drawing a sample from a population is a very fundamental concept in statistics. It is important to think of the objects or people in the population as *named* or *labeled*. For example, if there are n balls in a jar, we can imagine that they have labels from 1 to n, even if the balls look the same to the human eye. In the birthday problem, we can give each person an ID (identification) number, rather than thinking of the people as indistinguishable particles or a faceless mob.

Adjusting for overcounting

In many counting problems, it is not easy to directly count each possibility once and only once. If, however, we are able to count each possibility exactly c times for some c, then we can adjust by dividing by c. For example, if we have exactly double-counted each possibility, we can divide by 2 to get the correct count. We call this *adjusting for overcounting*.

Example 1.4.11 (Committees and teams).

Consider a group of four people.

(a) How many ways are there to choose a two-person committee?

(b) How many ways are there to break the people into two teams of two?

Solution

(a) One way to count the possibilities is by listing them out: labeling the people as 1, 2, 3, 4, the possibilities are $\boxed{1,2}$, $\boxed{1,3}$, $\boxed{1,4}$, $\boxed{2,3}$, $\boxed{2,4}$, $\boxed{3,4}$.

Another approach is to use the multiplication rule with an adjustment for overcounting. By the multiplication rule, there are 4 ways to choose the first person on the committee and 3 ways to choose the second person on the committee, but this counts each possibility twice, since picking 1 and 2 to be on the committee is the same as picking 2 and 1 to be on the committee. Since we have overcounted by a factor of 2, the number of possibilities is $(4 \cdot 3)/2 = 6$.

(b) Here are 3 ways to see that there are 3 ways to form the teams. Labeling the people as $1, 2, 3, 4$, we can directly list out the possibilities: $\boxed{1,2}\boxed{3,4}$, $\boxed{1,3}\boxed{2,4}$, and $\boxed{1,4}\boxed{2,3}$. Listing out all possibilities would quickly become tedious or infeasible with more people though. Another approach is to note that it suffices to specify person 1's teammate (and then the other team is determined). A third way is to use (a) to see that there are 6 ways to choose one team. This overcounts by a factor of 2, since picking 1 and 2 to be a team is equivalent to picking 3 and 4 to be a team. So again the answer is $6/2 = 3$.

A *binomial coefficient* counts the number of subsets of a certain size for a set, such as the number of ways to choose a committee of size k from a set of n people. Sets and subsets are by definition *unordered*, e.g., $\{3,1,4\}=\{4,1,3\}$, so we are counting the number of ways to choose k objects out of n, without replacement and without distinguishing between the different orders in which they could be chosen.

Definition 1.4.12 (Binomial coefficient).

For any nonnegative integers k and n, the *binomial coefficient* $\binom{n}{k}$, read as "n choose k", is the number of subsets of size k for a set of size n.

For example, $\binom{4}{2}=6$, as shown in Example 1.4.11. The binomial coefficient $\binom{n}{k}$ is sometimes called a *combination*, but we do not use that terminology here since "combination" is such a useful general-purpose word. Algebraically, binomial coefficients can be computed as follows.

Theorem 1.4.13 (Binomial coefficient formula).

For $k \leq n$, we have

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

For $k>n$, we have $\binom{n}{k}=0$.

Proof

Let A be a set with $|A|=n$. Any subset of A has size at most n, so $\binom{n}{k}=0$ for $k>n$. Now let $k \leq n$. By Theorem 1.4.6, there are $n(n-1)\cdots(n-k+1)$ ways to make an *ordered* choice of k elements without replacement. This overcounts each subset of interest by a factor of $k!$ (since we don't care how these elements are ordered), so we can get the correct count by dividing by $k!$.

Example 1.4.14 (Club officers).

In a club with n people, there are $n(n-1)(n-2)$ ways to choose a president, vice president, and treasurer, and there are $\binom{n}{3}=\frac{n(n-1)(n-2)}{3!}$ ways to choose 3 officers without predetermined titles.

Example 1.4.15 (Permutations of a word).

How many ways are there to permute the letters in the word LALALAAA? To determine a permutation, we just need to choose where the 5 A's go, in the 8 available slots (or, equivalently, choose where the 3 L's go). So there are

$$\binom{8}{5} = \binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{3!} = 56 \text{ permutations.}$$

We can use binomial coefficients to calculate probabilities in many problems for which the naive definition applies.

Example 1.4.16 (Full house in poker).

A 5-card hand is dealt from a standard, well-shuffled 52-card deck. The hand is called a *full house* in poker if it consists of three cards of some rank and two cards of another rank, e.g., three 7's and two 10's (in any order). What is the probability of a full house?

Solution

All of the $\binom{52}{5}$ possible hands are equally likely by symmetry, so the naive definition is applicable. To find the number of full house hands, use the multiplication rule (and imagine the tree). There are 13 choices for what rank we have three of; for concreteness, assume we have three 7's and focus on that branch of the tree. There are $\binom{4}{3}$ ways to choose which 7's we have. Then there are 12 choices for what rank we have two of, say 10's for concreteness, and $\binom{4}{2}$ ways to choose two 10's. Thus,

$$P(\text{full house}) = \frac{13\binom{4}{3}12\binom{4}{2}}{\binom{52}{5}} = \frac{3744}{2598960} \approx 0.00144.$$

The decimal approximation is more useful when playing poker, but the answer in terms of binomial coefficients is exact and *self-annotating* (seeing "$\binom{52}{5}$" is a much bigger hint of its origin than seeing "2598960").