# Lecture Notes

The last two lectures concerned the number of ways to choose elements from a finite set. In particular, we considered the number of ways to choose $r$ distinct elements where order does not matter ($r$-combinations) and where order does matter ($r$-permutations).

In this lecture we consider the number of ways to choose $r$ elements from a finite set where we are allowed to choose the same element repeatedly. Again, we look at both the case where order does not matter ($r$-combinations with repetition) and where order does matter ($r$-permutations with repetition).

## 1 Combinations with Repetition

Let $S$ be the set $\{A, B, C\}$. As we saw in the previous lecture, this set has three 2-combinations. That is, there are three ways to choose two distinct elements of $S$ where order does not matter. The three 2-combinations of $S$ are shown below.

$$\{A, B\} \quad \{A, C\} \quad \{B, C\}$$

Suppose that we are not required to choose distinct elements of $S$, but rather can choose the same element repeatedly. The resulting sets are called the *$r$-combinations with repetition* of the set $S$. Listed below are the six 2-combinations with repetition of $S$.

$$\{A, B\} \quad \{A, C\} \quad \{B, C\} \quad \{A, A\} \quad \{B, B\} \quad \{C, C\}$$

Strictly speaking, these are multisets, not sets, since an element may appear multiple times.

### 1.1 Counting $r$-Combinations with Repetition

The following theorem gives a nice formula for the number of $r$-combinations with repetition of an $n$-element set.

**Theorem 1.1** The number of $r$-combinations with repetition of an $n$-element set is

$$\binom{n+r-1}{r}.$$

In the example above, we found six ways to choose two elements from the set $S = \{A, B, C\}$ with repetition allowed. Sure enough, the theorem says that the number of 2-combinations of a 3-element set is $\binom{3+2-1}{2} = 6$.

For comparison, recall that the number of ordinary $r$-combinations of an $n$-element set is $\binom{n}{r}$. Every ordinary $r$-combination is also a valid $r$-combination with repetition. So, as one would expect, the number of $r$-combinations with repetition is greater if $r > 1$.

The proof of this theorem uses an important trick calls "stars and bars".

*Proof.* Let $S$ be a set with $n$ elements that are ordered in some way. We will establish a bijection between $r$-combinations with repetition of the set $S$ and strings of stars and bars.

Let $R$ be a particular $r$-combination with repetition of $S$. Write down $n-1$ bars. These $n-1$ bars divide the line into $n$ regions.

$$\underbrace{\quad \mid \quad \mid \quad \mid \quad \cdots \quad \mid \quad}_{n-1 \text{ bars define } n \text{ regions}}$$

Put one star in the $i$-th region for each time that the $i$-th element of $S$ appears in $R$. This procedure maps each $r$-combination with repetition to a string with $r$ stars and $n-1$ bars.

(For example, let $S$ be the set $\{A, B, C, D, E\}$, with elements ordered alphabetically. Let $R$ be the 7-combination with repetition $\{A, B, B, B, D, E, E\}$. The stars-and-bars string corresponding to $R$ is shown below.

$$\underbrace{\star}_{A} \mid \underbrace{\star\ \star\ \star}_{B,B,B} \mid \mid \underbrace{\star}_{D} \mid \underbrace{\star\ \star}_{E,E}$$

The two bars with no stars between indicate that element $C$ never appears in $R$.)

This mapping is a bijection because it has an inverse. That is, given any stars-and-bars string, we can construct the corresponding $r$-combination with repetition. The number of stars in the first region determines the number of times that the first element of $S$ appears in the $r$-combination, the stars in the second region determine the number of times that the second element appears, etc.

Since the mapping is a bijection, the number of $r$-combinations with repetition of an $n$-element set is equal to the number of strings containing $n-1$ bars and $r$ stars. The number of such strings is equal to the number of ways to choose $r$ distinct positions for the stars in a string of $n+r-1$ stars and bars. This is the number of ordinary $r$-combinations of a set with $n+r-1$ elements, which is $\binom{n+r-1}{r}$. ∎

## 1.2 Triple-Scoop Ice Cream Cones

Baskin-Robbins is an ice cream store that has 31 different flavors. How many different triple-scoop ice cream cones are possible at Baskin-Robbins? Two ice cream cones are considered the same if one can be obtained from the other by reordering the scoops. Of course, we are permitted to have two or even three scoops of the same flavor.

Of course, the best solution to this problem is to go to the Baskin-Robbins in Harvard Square and explicitly construct all possible combinations. This is called the *consumption method.* However, there is also a purely mathematical approach.

The number of triple-scoop ice cream cones is precisely the number of 3-combinations with repetition of the set of 31 flavors. Therefore, we can count the number of ice creams cones with the formula in Theorem 1.1 where $n = 31$ and $r = 3$. This gives:

$$\binom{31+3-1}{3} = \binom{33}{3} = \frac{33 \cdot 32 \cdot 31}{3 \cdot 2 \cdot 1} = 5456$$

On Thursdays the Harvard Square Baskin-Robbins is run by an irritating woman who refuses to serve a cone with two or three scoops of the same flavor. How many different triple-scoop ice cream cones are possible on Thursday?

Now we must select an ice cream cone by choosing 3 *distinct* flavors from the complete set of 31. Therefore, the number of different cones is the number of ordinary 3-combinations of a 31-element set. This is:

$$\binom{31}{3} = \frac{31 \cdot 30 \cdot 29}{3 \cdot 2 \cdot 1} = 4495$$

As we would expect, the number of combinations with repetition is greater than number number of combinations without repetition. Permitting scoops of the same flavor gives an extra $5456 - 4495 = 961$ options.

## 1.3 Dozens of Dunkin Donuts

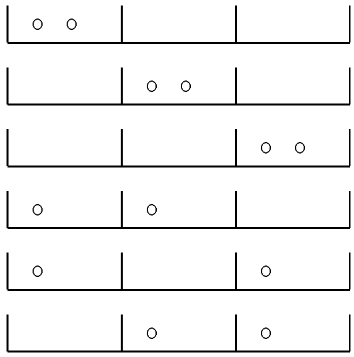
Suppose we next stagger down to the Dunkin Donuts in Central Square. We want a box of a dozen doughnuts and there are 21 varieties available. How many options do we have?

We select our box of doughnuts by pointing out 12 varieties from the complete set of 21 where repetition is allowed. Therefore, the number of options is the number of 12-combinations with repetition of a 21 element set. Applying Theorem 1.1 with $n = 21$ and $r = 12$, we find that number of different boxes of doughnuts is:

$$\binom{21+12-1}{12} = \binom{32}{12} = 225,792,840$$

## 1.4 Balls and Bins

Suppose that we have $r$ identical balls and $n$ distinct bins. In how many ways can we arrange the balls in the bins? There are six possibilities for the case of two balls and three bins, as shown below.



**Claim 1.2** There are $\binom{n+r-1}{r}$ ways to arrange $r$ identical balls in $n$ distinct bins.

*Proof.* Each arrangement of the $r$ balls corresponds to an $r$-combination with repetition of the set of $n$ bins. Specifically, if the $i$-th bin contains $a_i$ balls, then the corresponding $r$-combination with repetition contains the $i$-th bin $a_i$ times. Therefore, the number of arrangement of balls is equal to the number of $r$-combinations of an $n$ element set, which is $\binom{n+r-1}{r}$ by Theorem 1.1. ■

There is another way to prove the claim that uses a cute bijection between balls-and-bins arrangements and stars-and-bars strings. The six arrangements in the preceding example are redrawn below with some lines erased and the balls replaced by stars.

$$\star\ \star\ |\quad |$$
$$|\ \star\ \star\ |$$
$$|\quad |\ \star\ \star$$
$$\star\ |\ \star\ |$$
$$\star\ |\quad |\ \star$$
$$|\ \star\ |\ \star$$

Now each balls-and-bin diagram has become a stars-and-bars string. The number of ways to place $r$ balls in $n$ bins is therefore equal to the number of strings with $r$ stars and $n-1$ bars, which is $\binom{n+r-1}{r}$.

## 1.5 Data Compression

Stars and bars can help solve a data compression problem. How many bits are needed to specify a set of $N$ arbitrary integers in the range $[0, 2N]$? For example, how much disk space do we need to store one million numbers in the range zero to two million?

#### A Simple Scheme

One simple scheme is store each number in binary. With this approach, storing each number requires $\lceil \log_2(2N+1) \rceil$ bits. Storing all $N$ numbers requires $N \lceil \log_2(2N+1) \rceil$ bits. In the case of a million numbers, we need 21 bits to store each number and therefore 21 million bits to store all of the numbers.

#### A Better Approach Using Stars and Bars

There is a more clever scheme that uses stars and bars. We can regard the set of numbers that we want to store as an $N$-combination with repetition of the $(2N+1)$-element set $[0, 2N]$. The proof of Theorem 1.1 shows that such $N$-combinations with repetition correspond to strings of $N$ stars and $2N$ bars. If we represent a star with a 0 and a bar with a 1, then we can store such a stars-and-bars string using $N + 2N = 3N$ bits.

To make the scheme clear, suppose that we are storing a set of $N = 5$ numbers in the range $[0, 10]$. In particular, suppose that the numbers are $\{2, 4, 4, 6, 7\}$. We can represent this multiset as a stars-and-bars string and then with $3N = 15$ bits as follows.

$$\begin{aligned} \{2,4,4,6,7\} &\rightarrow \;|\;|\;\star\;|\;|\;\star\;\star\;|\;|\;\star\;|\;\star\;|\;|\;| \\ &\rightarrow \;1\;1\;0\;1\;1\;0\;0\;1\;1\;0\;1\;0\;1\;1\;1 \end{aligned}$$

As another example, we could store a million numbers in the range zero to two million using only 3 million bits, a factor of 7 improvement over the simple scheme. It is rather surprising that we need only 3 bits per number on average, no matter how large $N$ becomes!

#### An Optimal Scheme

The stars-and-bars approach is simple and efficient, but not quite optimal.

In any scheme, we must associate every set of $N$ numbers with a distinct binary string. From the stars-and-bars interpretation, we know that there are $\binom{3N}{N}$ possible sets of $N$ numbers. Therefore, we must be prepared to store any one of $\binom{3N}{N}$ distinct binary strings. Using $k$ bits, we can store at most $2^k$ different binary strings. This means that we need at least $k$ bits where $2^k \geq \binom{3N}{N}$. This inequality is satisfied provided $k \geq \left\lceil \log_2 \binom{3N}{N} \right\rceil$. Therefore, as a lower bound, we need at least $\left\lceil \log_2 \binom{3N}{N} \right\rceil$ bits.

How far from optimal is the stars-and-bars scheme? To answer this question, we must rewrite the lower bound in a more familiar form. To this end, we can use Stirling's Formula to approximate $\binom{3N}{N}$ as follows.

$$\begin{aligned}
\binom{3N}{N} &= \frac{(3N)!}{(2N)!\,N!} \\
&\sim \frac{\sqrt{2\pi(3N)}\left(\frac{3N}{e}\right)^{3N}}{\sqrt{2\pi(2N)}\left(\frac{2N}{e}\right)^{2N}\sqrt{2\pi N}\left(\frac{N}{e}\right)^{N}} \\
&= \sqrt{\frac{3}{4\pi N}}\left(\frac{3^{3N}}{2^{2N}}\right) \\
&= \sqrt{\frac{3}{4\pi N}}\left(\frac{27}{4}\right)^{N}
\end{aligned}$$

With this formula, we can approximate the lower bound.

$$\begin{aligned}
\left\lceil \log_2 \binom{3N}{N} \right\rceil &\sim N \log_2 \frac{27}{4} - \underbrace{\frac{1}{2}\log_2 \frac{4\pi N}{3}}_{\text{low-order term}} \\
&\sim 2.755\ldots \cdot N
\end{aligned}$$

The stars-and-bars approach uses $3N$ bits, which is slightly more than the lower bound of about $2.755N$ bits.

This lower bound is actually not hard to achieve, though we will not cover the details. The idea is to order all of the multisets of $N$ numbers in the range $[0, 2N]$ and then index them $0, 1, 2, \ldots$. We then store a multiset by storing its index. The index is a number in the range 0 to $\binom{3N}{N} - 1$, which we can store in exactly $\left\lceil \log_2 \binom{3N}{N} \right\rceil$ bits. Of course, the hard part is efficiently mapping a multiset to an index and vice versa. Nevertheless, we can store a million numbers in the range zero to two million with only 2.755 million bits.

### 1.6 $r$-Combinations with at Least One of Each Item

Suppose that Kaybee Toys carries balls in three delightful colors: red, blue, and green.[1] In how many ways can we choose five balls so that we get at least one ball of each color?

We might as well start by choosing one ball of each color. Then we can choose the last two balls however we like. Remember that we are counting combinations, so the order in which we choose the balls does not matter.

Under this interpretation, the number of options is equal to the number of ways to choose two balls from the set of three colors with repetition allowed. Therefore, by Theorem 1.1 there are $\binom{3+2-1}{2} = 6$ possibilities. Here they are:

[1] I bet you're wondering if the 6.042 staff has discovered a lucrative scam involving product promotion in lecture notes. No comment.

$$\begin{array}{lll} \{R,G,B,\ R,R\} & \{R,G,B,\ R,G\} & \{R,G,B,\ R,B\} \\ \{R,G,B,\ G,G\} & \{R,G,B,\ G,B\} & \{R,G,B,\ B,B\} \end{array}$$

This argument generalizes to give the theorem below.

**Theorem 1.3** The number of $r$-combinations with repetition of an $n$-element set that contain every element in the set at least once is:

$$\binom{r-1}{n-1}$$

For example, in the colored balls problem, we have $r = 5$ and $n = 3$. The theorem says that there are $\binom{5-1}{3-1} = 6$ possibilities, which is consistent with out previous answer.

*Proof.* Every such $r$-combination consists of the entire $n$-element set together with an $(r-n)$-combination with repetition of the $n$-element set. By Theorem 1.1 the number of such combinations is:

$$\binom{n+(r-n)-1}{r-n} = \binom{r-1}{r-n} = \binom{r-1}{n-1}$$

The last equality uses the identity $\binom{m}{k} = \binom{m}{m-k}$. ∎

How many $n$-combinations with repetition of an $n$-element set are there that contain every element in the set? In this case, Theorem 1.3 gives $\binom{n-1}{n-1} = 1$. This makes sense; the only such combination is the set $S$ itself!

In how many ways can we arrange $r$ identical balls in $n$ distinct bins so that no bin is empty? If we first put one ball in each bin, then we can arrange the remaining $r - n$ balls in $\binom{n+(r-n)-1}{r-n} = \binom{n-1}{r-1}$ ways.

## 1.7 $r$-Combinations with Limited Repetition

Suppose that we want to count $r$-combinations where an element can be repeated only a limited number of times. For example, in how many ways can we arrange 10 identical balls in 4 distinct bins such no bin gets more than 7 balls? A good way to solve such problems is first to count the number of arrangements without a limit on repetition and then to subtract off the illegal arrangements.

The number of ways to arrange 10 identical balls in 4 distinct bins without a limit on repetition is $\binom{4+10-1}{10} = 286$.

Now we must count the illegal arrangements; that is, arrangements in which some bin contains 8 or more balls. Since there are only 10 balls in total, at most one bin can be overloaded with 8 or more balls. We can count the number of ways to overload the first bin by putting 8 balls into the first bin and then observing that the last two balls can be placed in $\binom{4+2-1}{2} = 10$ ways. Since any one of the four bins could be overloaded, the total number of illegal arrangement is $4 \cdot 10 = 40$.

Overall, there are $286 - 40 = 246$ ways to arrange 10 balls in 4 bins so that no bin gets more than 7 balls.

Sometimes when there are limits on repetition, the number of illegal arrangements can itself be difficult to compute. In such cases, a messy inclusion-exclusion calculation may be necessary.

# 2 $r$-Permutations with Repetition

Now we move from combinations with repetition to permutations with repetition. An *$r$-permutation with repetition* of a set $S$ is the number of ways to choose $r$ elements from $S$ with repetition allowed and where order matters. For example, there are nine 2-permutations with repetition of the set $S = \{A, B, C\}$, which are listed below.

$$\begin{array}{ccc} (A, A) & (A, B) & (A, C) \\ (B, A) & (B, B) & (B, C) \\ (C, A) & (C, B) & (C, C) \end{array}$$

## 2.1 Counting $r$-Permutations with Repetition

Fortunately, $r$-permutations with repetition are easy to count.

**Theorem 2.1** The number of $r$-permutations with repetition of an $n$-element set is $n^r$.

For example, the theorem says that the number of 2-permutations with repetition of the 3-element set $S = \{A, B, C\}$ is $3^2 = 9$, which checks with our previous answer.

*Proof.* Let $S$ be a set with $n$ elements. The $r$-permutations with repetition of $S$ are precisely the elements of:

$$\underbrace{S \times S \times \ldots \times S}_{r \text{ terms}}$$

By the Product Rule, this set has cardinality $|S|^r = n^r$. ■

## 2.2 Permutations with Limited Repetition

We might want to count the number of $r$-permutations where each element can be repeated a limited number of times. In general, this leads to some hairy analysis and no closed-form answer. Therefore, we will consider only a special case, the number of $r$-permutations where each element is repeated a precisely specified number of times.

For example, in how many ways can we arrange the letters in the word $PEPPER$? This is equal to the number of 6-permutations of the set $\{P, E, R\}$ where $P$ is repeated 3 times, $E$ is repeated 2 times, and $R$ is repeated 1 time. (Since the total number of repetitions defines $r$, we will use the term "permutations" in place of "$r$-permutations" for the remainder of the section.)

Initially, suppose that we make all the letters distinct by adding subscripts. That is, we want the number of ways to arrange the letters $P_1E_1P_2P_3E_2R$. In this case, there are $6! = 720$ arrangements because there are six choices for the first letter, five choices for the second letter, etc.

Next, suppose that we erase the subscripts on the $E$'s. This maps each arrangement of the letters $P_1E_1P_2P_3E_2R$ to an arrangement of the letters $P_1EP_2P_3ER$. Since $E_1$ and $E_2$ could be ordered in 2! ways before the erasure, the mapping is 2!-to-1. For example, we have:

$$\begin{array}{ccc} P_1E_1P_2P_3E_2R & \rightarrow & P_1EP_2P_3ER \\ & \text{and} & \\ P_1E_2P_2P_3E_1R & \rightarrow & P_1EP_2P_3ER \end{array}$$

Therefore, by the Division Rule there are $\frac{6!}{2!} = 360$ arrangement of the letters in $P_1EP_2P_3ER$.

Finally, suppose that we erase the subscripts on the $P$'s. This maps each arrangement of the letters $P_1EP_2P_3ER$ to an arrangement of the letters $PEPPER$. Since $P_1$, $P_2$, and $P_3$ could be ordered in 3! ways before the erasure, this mapping is 3!-to-1. Therefore, by the Division Rule, the number of arrangements of the letters $PEPPER$ is $\frac{6!}{2!\ 3!} = 60$.

We can prove a general theorem using the same argument as in the $PEPPER$ problem.

**Theorem 2.2** Let $A$ be the set $\{a_1, a_2, \ldots, a_n\}$, and let $r_1, r_2, \ldots, r_n$ be non-negative integers. The number of permutations of the set $A$ where each element $a_i$ is repeated exactly $r_i$ times is:

$$\frac{(r_1 + r_2 + \ldots + r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}$$

For example, the theorem says that the number of permutations of $\{P, E, R\}$ where $P$ is repeated 3 times, $E$ is repeated 2 times, and $R$ is repeated 1 time is $\frac{(3+2+1)!}{3!\ 2!\ 1!} = 60$. This is the answer we found before.

*Proof.* *(Sketch.)* We initially make $r_i$ distinct copies of each element $a_i$. Then the number of permutations where each element is repeated exactly once is $(r_1 + r_2 + \ldots + r_n)!$. We then "erase the subscripts" on the distinct copies of element $a_1$. This defines an $r_1$!-to-1 mapping from old permutations to permutations where $a_1$ is repeated $r_1$ times and all other elements are repeated once. Therefore the number of new permutations is $\frac{(r_1+r_2+\ldots+r_n)!}{r_1!}$. Continuing this way with $a_2, a_3, \ldots$, we we find that the number of permutations where each element $a_i$ is repeated exactly $r_i$ times is $\frac{(r_1+r_2+\ldots+r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}$. ■

## 2.3 Multinomial Coefficients

What is the number of permutations of the set $A = \{a_1, a_2\}$ where $a_1$ is repeated $r_1$ times and $a_2$ is repeated $r_2$ times? According to Theorem 2.2, the answer is:

$$\frac{(r_1 + r_1)!}{r_1!\ r_2!} = \binom{r_1 + r_2}{r_1}$$

We can restate the question as follows. How many strings contain $r_1$ copies of the symbol $a_1$ and $r_2$ copies of the symbol $a_2$? This is equal to the number of ways to choose $r_1$ distinct positions for the $a_1$'s from the set of all $r_1 + r_2$ positions, which is $\binom{r_1+r_2}{r_1}$. This is the same as the previous answer.

By shifting from "permutations" to "strings" in this way, we can give an alternate proof of Theorem 2.2.

*Proof.* *(Alternate proof of Theorem 2.2.)* There is a bijection between permutations of the set $A = \{a_1, a_2, \ldots, a_n\}$ where each element $a_i$ appears exactly $r_i$ times and strings where each symbol $a_i$ appears exactly $r_i$ times. Therefore, we can count permutations by counting strings as follows.

The number of ways to choose $r_1$ distinct positions for the $a_1$'s from the set of all $r_1 + r_2 + \ldots + r_n$ positions is $\binom{r_1+r_2+\ldots+r_n}{r_1}$. Then the number of ways to choose $r_2$ distinct positions for the $a_2$'s from the set of all $r_2 + \ldots + r_n$ remaining positions is $\binom{r_2+\ldots+r_n}{r_2}$, and so forth. The total number of strings is therefore:

$$\begin{aligned}
& \binom{r_1 + r_2 + \ldots + r_n}{r_1} \cdot \binom{r_2 + \ldots + r_n}{r_2} \cdot \binom{r_3 + \ldots + r_n}{r_3} \cdots \binom{r_n}{r_n} \\
= \;& \frac{(r_1 + r_2 + \ldots + r_n)!}{r_1!\ (r_2 + \ldots + r_n)!} \cdot \frac{(r_2 + \ldots + r_n)!}{r_2!\ (r_3 + \ldots + r_n)!} \cdot \frac{(r_3 + \ldots + r_n)!}{r_3!\ (r_4 + \ldots + r_n)!} \cdots \frac{r_n!}{r_n!\ 0!} \\
= \;& \frac{(r_1 + r_2 + \ldots + r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}
\end{aligned}$$

The first equality uses the definition of binomial coefficients, and the second follows by cancelling terms. ∎

The quantity $\frac{(r_1+r_2+\ldots+r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}$ is called a *multinomial coefficient* and is denoted $\binom{r_1+r_2+\ldots+r_n}{r_1, r_2, \ldots, r_n}$. Given a set with $r_1 + r_2 + \ldots + r_n$ elements, the multinomial coefficient $\binom{r_1+r_2+\ldots+r_n}{r_1, r_2, \ldots, r_n}$ represents the number of ways to choose $r_1$ elements, then $r_2$ of the remaining elements, and so forth.

Multinomial coefficients also arise in the Multinomial Theorem, a generalization of the Binomial Theorem. The result is stated below, but not proved.

**Theorem 2.3 (Multinomial Theorem)**

$$(x_1 + x_2 + \ldots + x_n)^r = \sum^{r_1+r_2+\ldots+r_n=r} \binom{r}{r_1, r_2, \ldots, r_n} x^{r_1} x^{r_2} \ldots x^{r_n}$$

## 3 Summary

The table below summarizes the basic results from the last three lectures about choosing $r$ elements from an $n$-element set:

$$
\begin{aligned}
r\text{-combinations without repetition} &= \binom{n}{r} \\
r\text{-permutations without repetition} &= \frac{n!}{(n-r)!} \\
r\text{-combinations with repetition} &= \binom{n+r-1}{r} \\
r\text{-permutations with repetition} &= n^r
\end{aligned}
$$