1 Onto functions and bijections – Applications to Counting

Now we move on to a new topic.

Definition 1.1 (Surjection). *A function $f : A \to B$ is said to be surjective or onto if for each $b \in B$ there is some $a \in A$ so that $f(a) = B$.*

What are examples of a function that is surjective. That is not surjective?

Definition 1.2 (Bijection). *A function $f : A \to B$ is called a bijection if f is both one-to-one and surjective.*

Bijections are of particular interest to us. We have been talking informally about the size of certain sets. For instance, we have asked how many ways there are to list k elements from a set of n (either with or without repetition). This problem is the same as asking for the size of the set of such lists. Bijections can be thought of as a way of paring elements of two sets together. Each member of each set must be paired with exactly one member of the other set. Thus, intuitively, we want two sets to have the same size iff we can form a bijection between them. This leads to the following more formal definition of size:

With the concept of bijection, we can formally define the size of a set:

Definition 1.3 (Size). *A set A is finite if there is a bijection from A to $[n] = \{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$. In this case we say that A has size n.*

Fact 1.4. *A finite set must have exactly one size. That is, if $|A| = n$ and $|A| = m$, then $m = n$.*

Proof. Actually, the Exercise. Uses induction. □

Definition 1.5 (Composition). *Let $f : A \to B$ and $G : B \to C$. We say that $g \circ f : A \to C$ is the composition of g with f, and is defined by $g \circ f(a) = g(f(a))$.*

Fact 1.6. *If g and f are as defined above, then $g \circ f$ is indeed a function and furthermore:*

1. *If g and f are both onto, then so is $g \circ f$*
2. *If g and f are both one-to-one, then so is $g \circ f$*
3. *If g and g are both bijections, then so is $g \circ f$*

Proof. To see that $g \circ f$ is a function, let $a \in A$ be given. We must prove that there is a unique $c \in C$ so that $(a, c) \in g \circ f$. (We don't have the right to use the $g \circ f(a) = c$ notation until we prove that $g \circ f$ is a function). But we do know that since $f(a)$ is a unique value of b, that there is a unique c so that $g(f(a)) = c$. Therefore, by definition, there is exactly one c so that $(a, c) \in g \circ f$.

Now, I will prove only item (1). I leave items (2) and (3) as an exercise. So let g and f be surjections. Let $c \in C$ be given. Then since g is onto there is a $b \in B$ so that $g(b) = c$. And since f is onto there is an $a \in A$ so that $f(a) = b$. Therefore $g(f(a)) = g(b) = c$, so $g \circ f$ is onto. □

Fact 1.7. *Let f be a bijection from $A \to B$. Then the inverse relation of f, defined by $f^{-1} = \{(y, x) \mid (x, y) \in f\}$ is a function, and furthermore is a bijection.*

Proof. First we show that f^{-1} is a function from B to A. Suppose that $b \in B$. Then since f is a bijection, there is a unique $a \in A$ so that $f(a) = b$. Therefore, there is a unique $a \in A$ so that $(a, b) \in f$, so there is a unique $a \in A$ so that $(b, a) \in f^{-1}$. Therefore, f^{-1} is a function so that if $f(a) = b$ then $f^{-1}(b) = a$.

Now, we show that f^{-1} is a bijection. Here I will only show that f is one-to-one. I leave as an exercise the proof that f is onto. So let $f^{-1}(b_1) = f^{-1}(b_2) = a$ for some $b_1, b_2 \in B$ and $a \in A$. We need to show that $b_1 = b_2$. To see this, notice that since f is a function, and we know that $f^{-1}(b) = a$ iff $f(a) = b$, that $f(a) = b_1$ and $f(a) = b_2$. But this means that $b_1 = b_2$, because f is a function. □

Theorem 1.8. *Two sets A and B have the same size iff there is a bijection $f : A \to B$.*

Proof. Suppose that A and B both have size n. Then there is a bijection $g : [n] \to A$ and a bijection $h : [n] \to B$. Then g^{-1} is a bijection from A to $[n]$, so $h \circ g^{-1}$ is a bijection from A to B.

Now, suppose there is a bijection $f : A \to B$, and suppose that A has size n. Then there is a bijection $g : [n] \to A$. The function $f \circ g$ is a bijection from $[n]$ to B, so B must also have size n. □

2 Subsets

First, we will attempt to count the number of subsets of a set.

Definition 2.1 (Subset). *We say that $A \subseteq B$, or A is a <u>subset</u> of B, if each element of A is also an element of B. Note that this means that the set with no elements, called $\emptyset$, or the <u>empty set</u> is a subset of every set.*

One natural question is to ask, given a set A with n elements, how many subsets of A are there?

Fact 2.2. *The number of subsets of an n element set S is 2^n*

Proof. Label the n elements of the set $a_1, \ldots, a_n$. To determine for a subset A of S , for each i we just decide whether a_i is in the subset. We thus have a list of n choices, each choice being "yes" or "no". By the product principle, there are therefore 2^n such lists of choices. □

We can think of the above procedure in terms of functions. For each subset A of $[n]$, for instance, we define the function f_A by $f_A(x) = 1$ if $x \in A$, and $f_A(x) = 0$ if $x \notin A$. Thus, 1 corresponds to a "yes" in our list of choices and 0 to a "no".

Example Suppose that we have a class of 7 people and we want a (non empty) set of these 7 people to present a paper. How many possible groups are there? There are 2^7 possible subsets, but we are not allowing the empty set here, so the answer is $2^7 - 1$ or 127.

Now, suppose that there are no volunteers, and instead we have to choose a group of three people to present. How many ways can we do this? There are $7!/(7-3)! = 7 \cdot 6 \cdot 5$ such lists. But each group of three people can be listed in $3!$ ways. So there are a total of $\frac{7!}{(7-3)!3!} = 35$ different groups.

3 Binomial Coefficients

The above example suggests a more general question, and one that has widespread applications. Suppose we have a set of n people. We want to choose a set of k of them for some task. How many ways can we choose them?

Theorem 3.1. *If $0 \leq k \leq n$, then the number of k-element subsets of an n element set is given by the formula:*

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Proof. Let's consider the set of all lists of k elements choosen from a set of n elements. The number of ways to do this is $n!/(n-k)!$. Furthermore, we can generate all such lists by choosing k people first, and then listing them in the $k!$ proper ways. Thus there are $\binom{n}{k}k!$ ways of listing k people from a group of n. Therefore, $\binom{n}{k}k! = \frac{n!}{(n-k)!}$, so $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ □

Fact 3.2. $\binom{n}{k} = \binom{n}{n-k}$

Proof. Exercise. □

4 Using Pascal's Triangle

See the bottom of p.31 in Bogart, 3rd edition, for a pictoral view of Pascal's triangle. Pascal's triangle is contructed by placing a 1 at the top point, and a pair of 1s, one below and to the left and one beloe and to the right of the top 1. The further entries are computed by adding the ones above it. It turns out that if we count the rows starting at row 0, and the entries at each row from zero (So, the "first" row is now the 0th row, and so forth), then it turns out that $\binom{n}{k}$ is given by picking the kth element of the nth row. This fact is precisely given by the following theorem:

Fact 4.1. $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ *whenever* $0 < k < n$.

Proof. We consider the ways of picking k elements out of a set of n elements. Let S be a set of n elements. Let $s \in S$ be singled out. To pick k elements of S, we may either do so by choosing s, or not choosing s. If we elect to not choose

s, then we choose k of the remaining $n-1$ members, which can be done in $\binom{n-1}{k}$ waus. And if we elect to choose s, we choose only $k-1$ of the remaining $n-1$ members, which can be done in $\binom{n-1}{k-1}$ ways. So we have two disjoint collections of k-element subsets of S, which cover all of the possible k-element subsets of S. By the sum principle, we therefore get $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$. $\square$

Why do we care about Pascal's triangle? Think computationally. Suppose we want to compute $\binom{20}{10}$. If we use the formula as stated, we would have $\frac{20!}{10!10!}$. 20! and 10! are very large numbers. But when we evaluate the fraction, we get something more manageable: about 185,000. So this suggests that there may be a better way of calcluating $\binom{n}{k}$ – one where the numbers do not get very large. And this is Pascal's triangle. It turns out as we said above, we can easily get $\binom{n}{k}$ by looking at Pascal's triangle. And the entries in Pascal's triangle do not grow as quickly as factorials do, so the computational process is much easier.

So far, we have only defined $\binom{n}{k}$ for values where it makes sense to choose k objects out of n. But we can also define $\binom{n}{k}$ for other values. We do so by noticing that if $n \geq 0$, but $k < 0$ or $k > n$, then there *is no way* to choose k objects out of n. So in these cases we define $\binom{n}{k} = 0$. Later on, we will define $\binom{n}{k}$ when n is negative.

This method of using Pascal's Triangle to compute $\binom{n}{k}$ is an example of *recursion.* Recursion is when we try and compute a value for a function from values with smaller inputs. In other words, If we had a function f of one variable, and we had a way of computing $f(n)$ from $f(n-1)$.

5 Using Binomial Coefficients

Theorem 5.1 (Binomial theorem). *For any integer $n \geq 0$:*

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{o} x^i y^{n-i}$$

Proof. $(x+y)^n$ is a product of n factors. Each term in the factor is a result of picking either x or y from each of the n $(x+y)$ occurring in the term. $x^i y^j$ appears in the expansion only if $i+j=n$. Furthermore it appears in as many ways as we can choose exactly i x's out of the n factors. So the coefficient of $x^i y^{n-i}$ is $\binom{n}{i}$.

The result follows. $\square$

See examples 4.1 and 4.2 in Bogart.

Here's a more interesting example of this fact:

This is called Vandermonde's formula

Theorem 5.2. *Vandermonde's Formula is:*

$$\binom{m+n}{k} = \sum_{i=0}^{k} \binom{m}{i} \binom{n}{k-i}$$

Proof.

$$
\begin{aligned}
(x+1)^m(x+1)^n &= \sum_{i=0}^{m}\binom{m}{i}x^i\sum_{j=0}^{n}\binom{n}{j}x^j \\
&= \sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}x^i\binom{n}{j}x^j \\
&= \sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}x^i x^j \\
&= \sum_{k=0}^{m+n}\left(\sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}\right)x^k
\end{aligned}
$$

Since the coefficient of x^k is also $\binom{m}{i}\binom{n}{k-i}$, we get the desired formula. □

We've given a purely algebraic proof of the formula. Let's now give a more intuitive, combinatorial one:

Proof. The left hand side of the formula is $\binom{m+n}{k}$, choosing k elements out of a set of $m+n$. We can choose these elements by looking separately at the first m and the last n elements. For each i, we can get k elements of the large set of $m+n$ elements by choosing i of the first m and $k-i$ of the last n. By the product principle, there are $\binom{m}{i}\binom{n}{k-i}$ ways of doing this. Since each choice of i gives us a completely different collection of k elements, we get, from the sum principle:

$$\binom{m+n}{k} = \sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}$$

□