Counting

Lecturer: Peter Shor

In these notes we discuss techniques for counting classes of objects satisfying certain properties. This is useful both for evaluating the number of possibilities of a certain phenomenon, but also when dealing with probabilities, to calculate some potentially tricky probability. In particular, in these notes, we discuss basic counting tools like combinations/permutations, counting in two different ways, and its formalization to bijective proofs for counting. To illustrate the latter, we will look at many different counting problems whose answer is $C_n = \frac{1}{n+1}\binom{2n}{n}$, known as the nth Catalan number. (The name has nothing to do with Catalonia, but is the name of a Belgian mathematician, Eugène Charles Catalan, who discovered it first in the 19th century.)

In later notes, we'll introduce another technique for counting through the use of generating functions.

1 Introduction: combinations and permutations

Let's recall the formula

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

for choosing k things from n things. This is called the *binomial coefficient* and often pronounced n *choose* k.

How do you prove that this is the number of ways of choosing a set of k items from n items total? You can pick the first items in n ways, the second item in $n-1$ ways, the third item in $n-2$ ways, and so forth. This gives $n(n-1)(n-2)\dots(n-k+1)$ ways of picking k **ordered** sets out of n items.

However, we're trying to count unordered sets. Each unordered set of k elements corresponds to $k!$ different ordered sets, one for each *permutation* of the elements of the unordered set. For example, if we were trying to choose 2 numbers from $\{1, 2, \dots 7\}$, the unordered set $\{2, 5\}$ would correspond to the two ordered sets $(2, 5)$ and $(5, 2)$. We thus have to divide by $k!$ to obtain the number of unordered sets. This shows

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

Note that we can multiply the numerator and denominator on the right-hand side by $(n-k)!$ to obtain

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

Note that this expression is symmetric in the number of chosen items, k and the number of unchosen items, $n-k$. Thus, we see that $\binom{n}{k} = \binom{n}{n-k}$. We should have expected this, since we could also single out a set of size k by choosing $n-k$ items to remove.

A related thing we want to cover is the *multinomial coefficient*. Suppose we have n items, and we want to color them with j different colors. Suppose furthermore that we want to color exactly k_i objects with color i. Thus $n = \sum_{i=1}^{j} k_i$. The number of ways of doing this is the multinomial coefficient

$$\binom{n}{k_1\, k_2\, \dots k_j} = \frac{n!}{k_1!k_2!k_3! \dots k_j!}$$

One can prove this by using the binomial coefficient formula for choosing k items from n and induction on the number of colors. We won't write this up in detail, but it follows from the formula

$$\binom{n}{k_1\, k_2\, \dots\, k_j} = \binom{n}{k_j}\binom{n-k_j}{k_1\, k_2\, \dots\, k_{j-1}}.$$

Also, notice that by definition $\binom{n}{k\, n-k}$ is the binomial coefficient $\binom{n}{k}$.

2 Counting via counting in two different ways

The first technique we will see is how to count via interpreting a certain quantity in two different ways. Most often this is useful for proving certain combinatorial identities. Here are two classical examples.

Example (Pascal's identity): Let's prove

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

for all $1 \le k \le n$. The left hand side, as we saw, is the number of ways to choose a set of k items out of n items. Let's count this another way. Let's single out one of these elements, and call it a. Our k-element set can either include a or not. If it does, there are $\binom{n-1}{k-1}$ ways to complete the set: pick any set of $k-1$ items out the remaining $n-1$ items. If it does not, there are $\binom{n-1}{k}$ ways to complete the set: we need to pick a set of k items out the remaining $n-1$ items.

Example (Number of subsets): Let's prove the identity

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

Again, let's count things two different ways. The right hand side is equal to the number of ways to choose a subset (of any size) of a set of n items Indeed – any item can be either *included* or *excluded* in the subset, so there are 2^n ways of doing this. On the other hand, we can also for any $0 \le k \le n$, count the number of possible subsets of size k, and add these up. By the previous section, we know that the number of possible subsets of size k is $\binom{n}{k}$, which proves the identity.

Implicitly, we have used the fact that we can represent any subset out of $\{1, 2, \cdots, n\}$ by a binary vector in $\{0, 1\}^n$. This correspondence can be extremely helpful when reasoning about uniformly random subsets of $\{1, 2, \cdots, n\}$, as each element of the set will be present with probability 0.5 and all these events are mutually independent.

Example (Counting number of trees on n vertices): This is a much more sophisticated and non-trivial illustration of the technique. Cayley's theorem is a classical result in graph theory which says that the number of different trees (subgraphs that are connected and don't have a cycle) that can be formed on n labeled vertices is n^{n-2}. What does it mean that the vertices are labeled? Informally that the tree on 3 vertices given $1-2-3$ is not the same as the one $2-1-3$, but the same as $3-2-1$ as it connects precisely the same set of vertices. Figure **??** (taken from Mathworld) illustrates this:

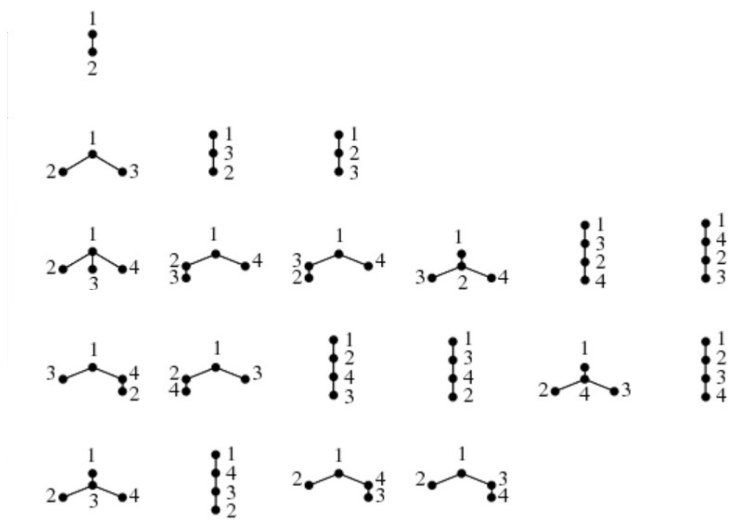

Figure 1: Trees on up to $n = 4$ labels.

For $n = 2$, we have only $2^0 = 1$ tree; for $n = 3$, we have $3^1 = 3$ trees (as we can remove any one of the 3 edges of a triangle), and so on.

Here we will give a beautiful proof of Cayley's theorem due to Jim Pitman by counting trees in two different ways.

For this purpose, let's introduce one more object: a rooted tree, which is a tree (for the purposes of this example, on labeled vertices) with a root, where the edges are directed towards their parents. (See Figure **??**). The *head* of each directed edge is the parent, and the *tail* is the child. The object we will count in two ways is the number of ordered sequences $(e_1, e_2, \ldots, e_{n-1})$ of edges that can be added to the empty set of labeled vertices in that order, so that the result is a proper rooted tree. Let's call this number L, and let's denote by T the number of different trees, which is what we are trying to count.

The first way to count will express this number via the number of trees. Towards this, note that a tree, along with a choice of root in it uniquely determines a rooted tree, since each of the edges must point from a node towards its parent. This means that $L = Tn(n-1)! = Tn!$. Indeed – for any tree, we pick a root (n choices), which determines the set of edges in it, but since we are counting ordered sequences $(e_1, e_2, \ldots, e_{n-1})$, we can order the $n-1$ edges in the tree in $(n-1)!$ different ways.

On the other hand, we will count L by reasoning in how many ways we can "add" the edges one-by-one to form a rooted tree. Suppose we have added k edges so far – we will reason about

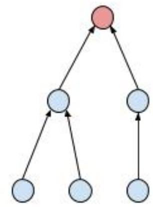

Figure 2: An example of a rooted tree

what E_k is – the number of possible choice for the next edge to add. Note that since we are forming a tree, after we have added k edges we will have a forest – a disjoint set of trees. Moreover, the number of trees in this forest will be $n - k$: indeed, any time we add an edge, we must merge two trees in the forest, otherwise we will form a loop; also, whenever we do this, we can never merge more than two trees. (See Figure **??**). So the endpoints of our edge must be in different

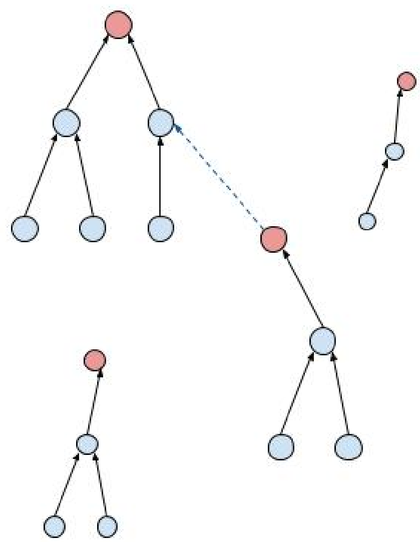

Figure 3: Adding new edge while building a rooted tree – dashed blue line is "added" edge.

components. We can choose the head of the edge arbitrarily (so there are n possible choices) but the tail should be one the roots in a different connected component (so there are $n-k-1$ choices). Therefore, $E_k = n(n-k-1)$. Altogether, we get

$$L = \Pi_{k=0}^{n-2} E_k = n^{n-1} \Pi_{k=0}^{n-2} (n-k-1) = n^{n-1}(n-1)!$$

So, by counting two ways, we get that $Tn! = L = n^{n-1}(n-1)!$, which implies that $T = n^{n-2}$.

3 Counting by bijections

A *bijection* is a one-to-one mapping from one set onto the other set, and if we have a bijection between two sets, we know that both sets have the same number of elements. This provides a nice technique for counting a collection of objects. One needs to find another collection of objects that is simpler to count, and construct a bijection between the two sets.

Before seeing examples, let us define what we mean by a bijection. Suppose we have a function f (also called a *map*) from a set A to a set B. This function f is a *bijection* if it is both an *injection* and a *surjection*. A function f mapping set A to set B is an *injection* if it doesn't map two different elements to the same element. In other words, if $f(x) = f(y)$, then $x = y$. A function f is a *surjection* if every element y of B has some element x of set A such that $f(x) = y$.

One way to prove that a map f is an injection is to show that it has an inverse. If there is a map g taking B to A such that $g(f(x)) = x$, then x must be an injection. If it is also true that $f(g(y)) = y$ for all $y \in B$, then it follows that f is also a surjection.

Let us consider some examples.

Counting colorings: Suppose you have n balls and j different colors. You can pick a color for each ball, and you want to know how many different sets of colored balls you can make. However, the balls are all identical so you only care about how many balls you have of each color.

For example, if you have three colors and two balls, the number of sets of colored balls you can make is six: RR, BB, GG, RB, RG, BG. (The order of the balls doesn't matter, so RG is the same as GR.) We will show that this is counted by the binomial coefficient $\binom{4}{2}$ by showing a bijection between ways of choosing two items out of four labeled items, and ways of choosing two balls of three colors.

How does this bijection work? Let's illustrate it by an example. Suppose we have three colors, and seven balls. Let's sort them by color in some canonical order, say R, G, B.

$$R\ R\ R\ R\ G\ B\ B$$

Now, let's put dividers between the colors

$$R\ R\ R\ R \mid G \mid B\ B$$

Now, let's forget the colors.

$$\mid \qquad \mid$$

We now have a sequence of $n + j - 1$ objects, n balls and $j - 1$ dividers. But the dividers tell us how to color the balls, so we can recover the information we started with, namely, how many balls were each color.

$$\mid \qquad \mid$$

This shows that the mapping (between n colored balls with j colors and the sequence of n uncolored balls and $j - 1$ dividers) is a bijection, and thus that the number of ways of coloring n balls with j colors is $\binom{n+j-1}{j-1}$.

Now, let's count one more thing before moving on to Catalan numbers. Suppose we have 30 objects and want to divide them into three sets of ten objects each. How many ways are there of doing this? We had

$$\binom{30}{10\,10\,10} = \frac{30!}{10!10!10!}$$

ways to color 30 objects with 10 red objects, 10 blue objects, and 10 green objects. If we want to count the ways of dividing them into uncolored sets, we need to divide by $6 = 3!$, because there are 6 ways of coloring three uncolored sets. A similar argument shows that the number of ways of partitioning $n = \sum_i k_i n_i$ objects into k_i sets of size n_i with the n_i's distinct is:

$$\frac{1}{\prod_i k_i!} \cdot \frac{n!}{\prod_i (n_i!)^{k_i}}.$$

The last case, when you have n unlabeled objects and you want to count the number of ways of dividing them into unlabeled sets (without specifying the size of the sets), is called the number of *partitions* of n. For example, fo $n = 5$, there are 7 partitions since we can write 5 as 5, as $4+1$, as $3+2$, as $3+1+1$, as $2+2+1$, as $2+1+1+1$ and as $1+1+1+1+1$. While there are lots of things you can say about partitions (including a nice generating function that counts them), there is no simple formula for the number of partitions.

4 Some Catalan families

We start by defining three classes of objects, and then discuss the relation between them.

A *plane tree* (a.k.a. ordered tree) is a rooted tree in which the order of the children matters. Let $\mathcal{T}_n$ be set of plane trees with n edges. The set $\mathcal{T}_3$ is represented in Figure **??**. A *binary tree* is a plane tree in which vertices have either 0 or 2 children. Vertices with 2 children are called *nodes*, while vertices with 0 children are called *leaves*. Let $\mathcal{B}_n$ be the set of binary trees with n nodes. The set $\mathcal{B}_3$ is represented in Figure **??**. A *Dyck path* is a *lattice path* (sequence of steps) made of steps +1 (up steps) and steps -1 (down steps) starting and ending at level 0 and remaining non-negative. Since the final level of a Dyck path is 0 the number of up steps and down steps are the same, and its length is even. Let $\mathcal{D}_n$ be the set of Dyck paths with $2n$ steps. The set $\mathcal{D}_3$ is represented in Figure **??**.

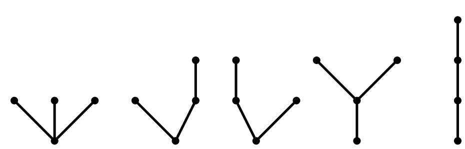

Figure 4: The set $\mathcal{T}_3$ of plane trees.

Observe that there is the same number of elements in $\mathcal{T}_3$, $\mathcal{B}_3$ and $\mathcal{D}_3$. This is no coincidence, as we will now prove that for all n, the sets $\mathcal{T}_n$, $\mathcal{B}_n$, $\mathcal{D}_n$ have the same number of elements. We now use the notation $|S|$ to denote the cardinality of a set S. We will now prove that

$$|\mathcal{T}_n| = |\mathcal{B}_n| = |\mathcal{D}_n| = \frac{1}{n+1}\binom{2n}{n}.$$

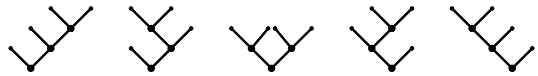

Figure 5: The set $\mathcal{B}_3$ of binary trees.

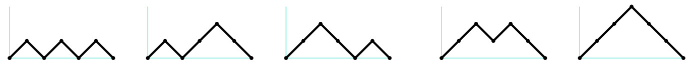

Figure 6: The set $\mathcal{D}_3$ of Dyck paths.

The number $\frac{1}{n+1}\binom{2n}{n}$ is the so-called nth *Catalan number*, and will be denoted by C_n. One can check that $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, and the Catalan sequence continues with $42, 132, 429, 1430, \cdots$. This is a beautiful sequence of numbers, which appears in many places in combinatorics.

Before proceeding, let us provide a probabilistic interpretation of this result (for Dyck paths). Suppose you have two candidates, Alice and Bob, for an election and there are $2n$ voters. Suppose there is a tie, and thus both candidates get precisely n votes in their favor. If you were to read the votes in a uniformly random ordering, what is the probability that Alice always has at least as many votes in her favor as Bob does, as one uncovers all the votes. You can easily check for yourself that this is equal to

$$\frac{|\mathcal{D}_n|}{\binom{2n}{n}} = \frac{C_n}{\binom{2n}{n}} = \frac{1}{n+1}.$$

For example, if you have 200 voters (and the resulting vote is a tie), this probability is 1/101.

4.1 Counting Dyck paths

We first compute the number of Dyck paths. Let $\mathcal{P}_n^{(0)}$ be the set of paths of length $2n$ made of steps $+1$ steps and -1 steps starting and ending at level 0. Ending at level 0 is the same as having the same number of up steps and down steps, and any choice of order of such steps is allowed. Hence

$$\mathcal{P}_n^{(0)} = \binom{2n}{n}.$$

Now $\mathcal{D}_n$ is a subset of $\mathcal{P}_n^{(0)}$. It seems hard to find $|\mathcal{D}_n|$ because of the non-negativity constraint, but actually a trick will now allow us to compute the cardinality of the complement subset

$$\overline{\mathcal{D}}_n \equiv \mathcal{P}_n^{(0)} \setminus \mathcal{D}_n.$$

Indeed we claim that $|\overline{\mathcal{D}}_n| = \binom{2n}{n-1}$. To prove this claim we consider the set $\mathcal{P}_n^{(-2)}$ of paths of length $2n$ made of steps $+1$ steps and -1 steps starting at level 0 and ending at level -2. These paths have $n-1$ up steps and $n+1$ down steps, and any order of steps is possible, hence $|\mathcal{P}_n^{(-2)}| = \binom{2n}{n-1}$. So it suffices to give a bijection f between $\overline{\mathcal{D}}_n$ and $\mathcal{P}_n^{(-2)}$. This bijection is defined as follows: take a path D in $\overline{\mathcal{D}}_n$ consider the first time t it reaches level -1. The path

$f(D)$ is obtained from D by flipping all the steps after time t with respect to the line $y = -1$. An example is shown in Figure **??**. We let the reader check that f is a bijection between $\overline{\mathcal{D}}_n$ and $\mathcal{P}_n^{(-2)}$. (Do it! For example, explicitly give the inverse g of f, show that $g(f(D)) = D$ for all $D \in \overline{\mathcal{D}}_n$ and $f(g(P)) = P$ for all $P \in \mathcal{P}_n^{(-2)}$.) Since f is a bijection we have $|\overline{\mathcal{D}}_n| = |\mathcal{P}_n^{(-2)}| = \binom{2n}{n-1}$.

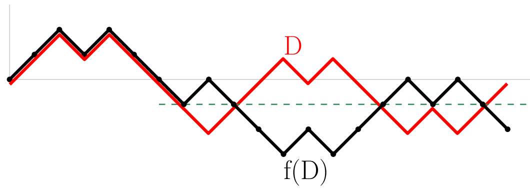

Figure 7: The bijection f: the path $D \in \overline{\mathcal{D}}$ in red, the path $f(D) \in \mathcal{P}_n^{(-2)}$ in black.

By the preceding, we have

$$|\mathcal{D}_n| = |\mathcal{P}_n^{(0)}| - |\overline{\mathcal{D}}_n| = \binom{2n}{n} - \binom{2n}{n-1} = \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n+1)!(n-1)!}.$$

And by reducing to the same denominator we find

$$|\mathcal{D}_n| = \frac{(2n)!}{n+1!n!} = \frac{1}{n+1}\binom{2n}{n},$$

as wanted.

4.2 Bijection between plane trees, binary trees and Dyck paths

We now present bijections between the sets $\mathcal{T}_n$, $\mathcal{B}_n$ and $\mathcal{D}_n$.

We first present a bijection Φ between plane trees and Dyck paths as follows: given any tree T in $\mathcal{T}_n$, perform a *depth-first search* of the tree T (as illustrated in Figure **??**) and define $\Phi(T)$ as the sequence of up and down steps performed during the search. A Dyke path is obtained from T because $\Phi(T)$ has n up steps and n down steps (one step in each direction for each edge of T), starts and end at level 0 and remains non-negative. A map Λ from Dyck paths to plane trees can also be easily defined in such a way that Φ and Λ are inverses of each other. Because Φ is a bijection between $\mathcal{T}_n$ and $\mathcal{D}_n$, we conclude

$$|\mathcal{T}_n| = |\mathcal{D}_n| = \frac{1}{n+1}\binom{2n}{n}.$$

We now present a bijection Ψ between binary trees and Dyck paths. Let B be a binary tree in $\mathcal{B}_n$. The tree B has n nodes. It can be shown that it has $n+1$ leaves (do it!). We can perform a depth-first search of the tree B and make a up step the first time we encounter each node and a down step each time we encounter a leaf. This makes a path with n up steps and $n+1$ down steps. The last step is a down step and we ignore it. We denote by $\Psi(B)$ the sequence of n up steps and n down steps obtained in this way. An example is represented in Figure **??**. It is actually true that

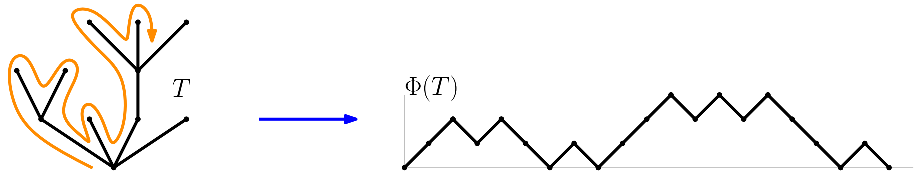

Figure 8: A plane tree T and the associated Dyck path $\Phi(T)$. The depth-first search of the tree T is represented graphically by a tour around the tree (drawn in orange).

$\Psi(B)$ is always a Dyck path and that Ψ is a bijection between $\mathcal{B}_n$ and $\mathcal{D}_n$. We omit the proof of these results. Since the sets $\mathcal{B}_n$ and $\mathcal{D}_n$ are in bijection we conclude

$$|\mathcal{B}_n| = |\mathcal{D}_n| = \frac{1}{n+1}\binom{2n}{n}.$$

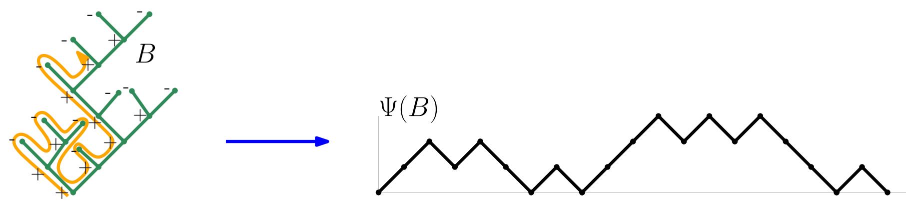

Figure 9: A binary tree B and the associated Dyck path $\Psi(B)$. The depth-first search of the tree B is represented graphically by a tour around the tree (drawn in orange).

5 Aside: Stirling's formula

Having a good approximation to the factorial is paramount in order to be able to know the order of growth of expressions involving them. The trivial bound that $n! \leq n^n$ and $n! \geq n \cdot (n-1) \cdot \dots \cdot n/2 \geq (n/2)^{n/2}$ immediately implies that $n \log n \geq \log(n!) \geq \frac{n}{2}\log(n/2)$ (where the log is in any base). Thus, aside for constants, $\log(n!)$ grows like $n \log n$. To have a more precise dependence, we can use Stirling's formula:

Lemma 1 (Stirling's formula)**.**

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

where $a \sim b$ means that $\frac{a}{b}$ tends to 1 as n tends to ∞.

Taking logarithm on both sides (in any base), we get that

$$\log(n!) = n\log(n) - n\log(e) + o(n), \tag{1}$$

where the "little o" notation means that the expression divided by n goes to zero as n goes to infinity. For natural logarithms, the expression simplifies to $\ln(n!) = n\ln(n) + o(n)$.

To gain familiarity with Stirling's formula, let's evaluate the middle binomial coefficient $\binom{2n}{n}$. By Stirling's formula, we have

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} \sim \frac{\sqrt{4\pi n}}{\sqrt{2\pi n}\sqrt{2\pi n}} \frac{(2n/e)^{2n}}{(n/e)^n(n/e)^n} = \frac{1}{\sqrt{\pi n}} 2^{2n} = \frac{1}{\sqrt{\pi n}} 4^n.$$

The fact that it is proportional to $4^n/\sqrt{n}$ has a probabilistic interpretation. Indeed, suppose we toss $2n$ fair coins, and want to evaluate the probability that we get exactly n heads (and n tails). This probability is equal to $\binom{2n}{n}/4^n$. Once we learn about Chebyshev's inequality and the weak law of large numbers, we will know that the fraction of heads is highly concentrated in an interval $(\frac{1}{2} - \frac{c}{\sqrt{n}}, \frac{1}{2} + \frac{c}{\sqrt{n}})$ for some constant c, and thus we expect the probability of getting precisely a fraction of half to be proportional to $1/\sqrt{n}$.

We will not derive Stirling's formula, but the following weakening for the middle binomial coefficient can be proved easily by induction on n:

$$\frac{4^n}{\sqrt{4n}} \leq \binom{2n}{n} \leq \frac{4^n}{\sqrt{3n+1}}.$$

This inductive proof is left as an exercise. In the base case of $n = 1$, we have equality throughout.

From Stirling's formula, we can derive that the nth Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$ satisfies

$$C_n \sim \frac{1}{\sqrt{\pi} n^{3/2}} 4^n.$$