Chapter 4

Binomial Coefficients

Binomial coefficients are seemingly somehow involved with almost every combinatorial construction.

4.1 BINOMIAL COEFF IDENTITIES

Table 4.1.1 **Pascal's triangle of binomial coefficients.**

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	Σ
0	1									1
1	1	1								2
2	1	2	1							4
3	1	3	3	1						8
4	1	4	6	4	1					16
5	1	5	10	10	5	1				32
6	1	6	15	20	15	6	1			64
7	1	7	21	35	35	21	7	1		128
8	1	8	28	56	70	56	28	8	1	256

REVIEW FROM §1.3 AND §1.5: We recall the following definitions and results.

- **Prop 1.3.1.** The combination coefficients $\binom{n}{k}$ satisfy ***Pascal's recurrence***:

$$\binom{n}{0} = 1 \qquad \text{for all } n \geq 0 \qquad \text{left column}$$

$$\binom{0}{k} = 0 \qquad \text{for all } k \geq 1 \qquad \text{top row}$$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \qquad \text{for } n \geq 1 \qquad (4.1.1)$$

- The ***binomial coefficient*** $b_{n,k}$ is the coefficient of x^k in the binomial expansion

$$(1+x)^n = \sum_{k=0}^{n} b_{n,k}\, x^k$$

- **Prop 1.3.2.** The binomial coeffs satisfy Pascal's recurrence.

- **Cor 1.3.3.** For all $k, n \in \mathbb{N}$,

$$\binom{n}{k} = b_{n,k}$$

- Since combination coeffs have exactly the same values as binomial coeffs, as per Corollary 1.3.3, they are commonly referred to as *binomial coefficients.*

- **Prop 1.5.3.** For all non-negative integers n and k,

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} \tag{4.1.2a}$$

- **Cor 1.5.4.** For all non-negative integers n and k,

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \tag{4.1.2b}$$

Combinatorial vs. Algebraic Proofs

TERMINOLOGY: An ***algebraic proof of an equation*** is achieved by transforming one side of the equation with the aid of substitutions and of arithmetic operations into the expression on the other side.

TERMINOLOGY: A ***combinatorial proof of an equation*** is achieved by showing that both sides of the equation count the same thing. Sometimes such a proof uses the pigeonhole principle.

Numerous examples of both kinds of proof follow. Sometimes we give two or more proofs of a single assertion. Various general methods, including mathematical induction, may be used with either type of proof.

Symmetry

Some identities are generalizations of properties readily noticeable in Pascal's triangle. One such property is that each row of Pascal's triangle is palindromic: it reads the same forward or backward. For instance, we observe the symmetry of row 8.

$$1 \quad 8 \quad 28 \quad 56 \quad 70 \quad 56 \quad 28 \quad 8 \quad 1$$

Proposition 4.1.1 [Row-Symmetry Property]. *For any integers n and k such that $0 \le k \le n$,*

$$\binom{n}{k} = \binom{n}{n-k} \tag{4.1.3}$$

Algebraic Proof: Using Eq. (4.1.2b) yields this easy algebraic proof.

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

Combinatorial Proof: The left side of Eq. (4.1.3) is the # ways to select k objects from the set of n, to be in the designated subset. The right side is the # ways to select $n-k$ objects to be excluded from it. There must be the same number of ways to do either. $\diamondsuit$

Row-Sum Property

Another property of Pascal's triangle is that the sum of the entries in each row is a power of 2. E.g., in row 8,

$$\begin{aligned} &1 + 8 + 28 + 56 + 70 + 56 + 28 + 8 + 1 \\ &\quad = 256 = 2^8 \end{aligned}$$

Proposition 4.1.2 [*Row-Sum Property*]. *The sum of the entries in row n of Pascal's triangle is 2^n. That is,*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \tag{4.1.4}$$

Combinatorial Proof: According to Corollary 1.3.3, the summands on the left side are the number of ways to choose subsets of cardinality k from a set S of n objects, for respective values of k. Their total is the number of ways to select a subset from S, over all possible subset sizes, which is clearly 2^n, shown on the right side, since each of the n objects is either present or absent. ♢

Algebraic Proof: Substituting $x = 1$ into both sides of the equation for a binomial expansion yields the following result.

$$\begin{aligned} (1+x)^n &= \sum_{k=0}^{n} \binom{n}{k} x^n \\ \Rightarrow \quad (1+x)^n \Big|_{x=1} &= \sum_{k=0}^{n} \binom{n}{k} x^n \Bigg|_{x=1} = \sum_{k=0}^{n} \binom{n}{k} 1^n \\ \Rightarrow \quad 2^n &= \sum_{k=0}^{n} \binom{n}{k} \end{aligned}$$ ♢

Algebraic Proof: Another algebraic proof is by induction, starting with row 0 of Pascal's triangle as a basis case, and then using Pascal's recursion to show that the sum in row n doubles the sum in row $n-1$. ♢

Column-Sum Property

Some other properties of Pascal's triangle emerge after further investigation. For instance, the sum of all the entries in any column, up to and including the entry in row n, can be found in the next column in row $n+1$.

Example 4.1.1: In columns 2 and 3 of Pascal's triangle, we see the following configuration.

n	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$
2		1	
3		3	
4		6	
5		10	
6		15	
7			$\boxed{35}$

$$1+3+6+10+15 \;=\; 35$$

Prop 4.1.3 [*Column-Sum Property*]. *The sum of the entries in column c $(c \geq 0)$ of Pascal's $\triangle$, from row 0 down to row n, equals the entry in row $n+1$, column $c+1$. I.e.,*

$$\sum_{k=0}^{n} \binom{k}{c} \;=\; \binom{n+1}{c+1} \qquad (4.1.5)$$

Proof: By induction on the row number n.

BASIS: For $n = 0$, the sum of the entries down to row 0 is 1, in column $c = 0$, and is otherwise 0; also,

$$\binom{1}{c+1} \;=\; \begin{cases} 1 & \text{if } c = 0 \\ 0 & \text{otherwise} \end{cases}$$

IND HYP: Assume for some $n \geq 1$ that

$$\sum_{k=0}^{n-1} \binom{k}{c} = \binom{n}{c+1}$$

IND STEP: Then

$$\begin{aligned}
\sum_{k=0}^{n} \binom{k}{c} &= \sum_{k=0}^{n-1} \binom{k}{c} + \binom{n}{c} \\
&= \binom{n}{c+1} + \binom{n}{c} && \text{(induction hypothesis)} \\
&= \binom{n+1}{c+1} && \text{(Pascal's recursion)} \; \Diamond
\end{aligned}$$

Diagonal-Sum Properties

DEF: A diagonal from the upper left of a 2-dimensional array, toward the lower right, is called a ***southeast diagonal***. A diagonal in the opposite direction is called a ***northwest diagonal***.

DEF: A diagonal from the lower left of a 2-dimensional array, toward the upper right, is called a ***northeast diagonal***. A diagonal in the opposite direction is called a ***southwest diagonal***.

We observe that the sum of the elements along a finite initial segment of the southeast diagonal in Pascal's triangle appears just below the southeasternmost entry.

Example 4.1.2: Here is a southeast diagonal-sum.

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$
2	1				
3		3			
4			6		
5				10	
6					15
7					$\boxed{35}$

$$1+3+6+10+15 \;=\; 35$$

Prop 4.1.4 [SE-Diagonal-Sum Property]. *The sum of the first $n+1$ entries on the southeast diagonal from row r, column 0 in Pascal's triangle equals the entry in row $r+n+1$, column n, the entry immediately below the last entry of the diagonal. That is,*

$$\sum_{k=0}^{n} \binom{r+k}{k} \;=\; \binom{r+n+1}{n} \tag{4.1.6a}$$

Proof: This result follows from two previously derived properties of Pascal's triangle.

$$\sum_{k=0}^{n} \binom{r+k}{k} \;=\; \sum_{k=0}^{n} \binom{r+k}{r} \qquad \text{(symmetry)}$$

$$= \binom{r+n+1}{r+1} \qquad \text{(column-sum property)}$$
$$= \binom{r+n+1}{n} \qquad \text{(symmetry)} \qquad \diamondsuit$$

The following corollary simply reverses the order of summation of the elements on the diagonal.

Cor 4.1.5 [NW-Diagonal-Sum Property]. *For any non-negative integer m such that $0 \leq m \leq n$, the binomial coefficients satisfy the equation*

$$\sum_{k=0}^{m} \binom{n-k}{m-k} = \binom{n+1}{m} \tag{4.1.6b}$$

Proof: Reversing the NW diagonal sum

$$\binom{n-0}{m-0} + \binom{n-1}{m-1} + \cdots + \binom{n-m}{m-m}$$

on the left of the equation yields the SE diagonal sum

$$\binom{n-m}{0} + \binom{n-m+1}{1} + \cdots + \binom{n}{m}$$

which starts at row $n-m$ and includes $m+1$ entries downward, ending at row n, column m. By Prop 4.1.4, the value of this southeast diagonal sum is the binomial coefficient

$$\binom{n+1}{m} \qquad \diamondsuit$$

The sums of northeast diagonals are Fibonacci numbers. For instance, the sum $1+5+6+1$ along the northeast diagonal that starts at $\binom{6}{0}$ is the Fibonacci number $f_7 = 13$.

Example 4.1.3: The boxed Fibonacci numbers shown here do not actually appear at the locations shown. They are simply the sums along the northeast diagonals that lead to them.

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$
1				$\boxed{5}$	
2			1		
3		3			$\boxed{21}$
4	1			4	
5			10		
6		6			
7	1				

$$1+3+1 = \boxed{5} = f_5$$

$$1+6+10+4 = \boxed{21} = f_8$$

Prop 4.1.6 [NE-Diagonal Fibonacci Property]. *The sum of the entries on the NE diagonal from row n, column 0 in Pascal's triangle equals the Fibonacci number f_{n+1}. That is,*

$$\sum_{k=0}^{n} \binom{n-k}{k} = f_{n+1} \tag{4.1.7}$$

Proof: BASIS: For $n = 0$ and $n = 1$ the northeast diagonal sums are $1 = f_1$ and $1 + 0 = f_2$, respectively.

IND HYP: For $n \geq 2$ assume that

$$\sum_{k=0}^{n-1} \binom{n-1-k}{k} = f_n \quad \text{and} \quad \sum_{k=0}^{n-2} \binom{n-2-k}{k} = f_{n-1}$$

IND STEP: By the Pascal recursion, we have

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

Therefore,

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n-k}{k} &= \sum_{k=0}^{n} \binom{n-k-1}{k-1} + \sum_{k=0}^{n} \binom{n-k-1}{k} \\
&= \sum_{j=0}^{n-2} \binom{n-j-2}{j} + \sum_{k=0}^{n-1} \binom{n-k-1}{k} \\
&= f_{n-1} + f_n \qquad \text{(induction hypothesis)} \\
&= f_{n+1} \qquad \text{(Fibonacci recursion)}\diamondsuit
\end{aligned}$$

Products of Binomial Coefficients

Another pattern in Pascal's triangle is the relationship between each element and the element to its upper left.

Example 4.1.4: We observe in this inset from Pascal's triangle that

$$4\binom{7}{4} \;=\; 4 \cdot 35 \;=\; 140 \;=\; 7 \cdot 20 \;=\; 7\binom{6}{3}$$

n	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$
5				
6		20		
7			35	
8				

$$\binom{7}{4} \cdot \frac{4}{7} \;=\; 35 \cdot \frac{4}{7} \;=\; 20 \;=\; \binom{6}{3}$$

or, equivalently

$$\binom{7}{4} \cdot 4 \;=\; 140 \;=\; \binom{6}{3} \cdot 7$$

The generality of this relationship, which is called the *absorption property*, is established by the next proposition.

Prop 4.1.7 [*Absorption Property*]. *For* $0 \le k \le n$,

$$\binom{n}{k} k \;=\; n\binom{n-1}{k-1} \tag{4.1.8}$$

Algebraic Proof: By algebraic manipulation, we have

$$\binom{n}{k} k \;=\; \frac{n^{\underline{k}}}{k!} k \;=\; \frac{n^{\underline{k}}}{(k-1)!} \;=\; n\frac{(n-1)^{\underline{k-1}}}{(k-1)!} \;=\; n\binom{n-1}{k-1}$$

$\diamondsuit$

Combinatorial Proof: Alternatively, we observe that the left side

$$\binom{n}{k} k$$

is the number of ways of choosing a board of k directors from a set of n persons and then a chairperson from within that board of k. This is clearly equivalent to the number of ways to choose a chairperson from a set of n persons and then another $k-1$ persons from the remaining $n-1$ persons for the rest of the board of directors, which is the right side

$$n\binom{n-1}{k-1} \qquad \diamondsuit$$

Absorption is a special case of a relationship between an element and other elements along the northwest diagonal direction. This relationship is expressed by a highly useful combinatorial identity that generalizes the following illustration.

Example 4.1.4, cont.: Observe that whereas at one position northwest of the coefficient $\binom{7}{4}$ we have

$$20 = \binom{6}{3} = \binom{7}{4}\frac{4^{\underline{1}}}{7^{\underline{1}}} = 35 \cdot \frac{4}{7}$$

at three positions northwest of $\binom{7}{4}$ in Pascal's triangle, we have

$$4 = \binom{4}{1} = \binom{7}{4}\frac{4^{\underline{3}}}{7^{\underline{3}}} = 35 \cdot \frac{24}{210}$$

n	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$
4	4				
5					
6					
7				35	
8					

$$35 \cdot \frac{4^{\underline{3}}}{7^{\underline{3}}} = \binom{7}{4} \cdot \frac{4^{\underline{3}} \,/\, 3!}{7^{\underline{3}} \,/\, 3!} = \binom{4}{1} = 4$$

or, equivalently

$$\binom{7}{4} \cdot \binom{4}{3} = \binom{7-3}{4-3} \cdot \binom{7}{3}$$

The following formulation of this property is called the *subset-of-a-subset property.*

Prop 4.1.8 [Subset-of-a-Subset Identity].
For $0 \le k \le m \le n$*,*

$$\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k} \tag{4.1.9}$$

Algebraic Proof: By straightforward algebraic calculation, we have

$$\begin{aligned}
\binom{n}{m}\binom{m}{k} &= \frac{n!}{m!\,(n-m)!} \cdot \frac{m!}{k!\,(m-k)!} \\
&= \frac{n!}{k!\,(n-m)!\,(m-k)!} \\
&= \frac{n!}{k!\,(n-k)!} \cdot \frac{(n-k)!}{(n-m)!\,(m-k)!} \\
&= \binom{n}{k}\binom{n-r}{m-k}
\end{aligned}$$

$\diamondsuit$

Combinatorial Proof: We can also reason combinatorially that the left side

$$\binom{n}{m}\binom{m}{k}$$

is the number of ways of choosing a board of m directors from a set of n persons and then an executive committee of k persons from within that board of m. This is clearly equivalent to the number of ways to choose an executive committee of k persons from a set of n persons and then another $m-k$ persons from the remaining $n-k$ persons for the rest of the board of directors, which is the right side

$$\binom{n}{k}\binom{n-k}{m-k} \qquad \diamondsuit$$

Vandermonde Convolution

Thm 4.1.9 [*Vandermonde Convolution*]. *Let n, m, and k be non-negative integers. Then*

$$\sum_{j=0}^{n}\binom{n}{j}\binom{m}{k-j} = \binom{n+m}{k} \tag{4.1.10}$$

Combinatorial Proof: A combinatorial proof supposes that there are $n+m$ objects in a set, n of them blue and m of them red, and that k objects are to be chosen, for which there are clearly $\binom{n+m}{k}$ ways in all, the number of the right side. The number of ways to select j blue objects and $k-j$ red objects is the product $\binom{n}{j}\binom{m}{k-j}$; so the sum of all these products, which is on the left side, must be the same total as the right side. $\diamondsuit$

Another Proof: The sum on the left of the combinatorial equation above equals the coefficient of x^k on the left side of the polynomial equation

$$(1+x)^n(1+x)^m \;=\; (1+x)^{n+m}$$

and the binomial coefficient on the right side of the combinatorial equation equals the coefficient of x^k on the right side of that polynomial equation. $\diamondsuit$

Summary of Binomial Coeff Identities

Table 4.1.2 Basic Binomial Coefficient Identities

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \quad \text{Pascal Rec} \quad (4.1.1)$$

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} \quad \text{Falling Power Formula} \quad (4.1.2a)$$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \quad \text{Factorial Formula} \quad (4.1.2b)$$

$$\binom{n}{k} = \binom{n}{n-k} \quad \text{Symmetry} \quad (4.1.3)$$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \qquad \text{Row-Sum} \quad (4.1.4)$$

$$\sum_{r=0}^{n} \binom{r}{c} = \binom{n+1}{c+1} \qquad \text{Column-Sum} \quad (4.1.5)$$

$$\sum_{k=0}^{n} \binom{r+k}{k} = \binom{r+n+1}{n} \qquad \text{SE Diagonal} \quad (4.1.6a)$$

$$\sum_{k=0}^{m} \binom{n-k}{m-k} = \binom{n+1}{m} \qquad \text{NW Diagonal} \quad (4.1.6b)$$

$$\sum_{k=0}^{n} \binom{n-k}{k} = f_{n+1} \qquad \text{Fibonacci NE Diagonal} \quad (4.1.7)$$

$$\binom{n}{k} k = n\binom{n-1}{k-1} \qquad \text{Absorption} \quad (4.1.8)$$

$$\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k} \qquad \text{Subset-of-a-Subset} \quad (4.1.9)$$

$$\sum_{j=0}^{n} \binom{n}{j}\binom{m}{k-j} = \binom{n+m}{k} \qquad \text{Vandermonde Convo} \quad (4.1.10)$$

Parity of Binomial Coefficients

Beyond the basics of binomial coefficients, there are many fascinating byways. For instance, how might one determine the parity of a given binomial coefficient, such as

$$\binom{165}{93}$$

without doing a lot of calculation? Scanning Pascal's triangle enhances the mystery. One observes that all the entries in rows 1, 3, and 7, numbers of the form $2^n - 1$, are odd. Moreover, the number of odd numbers in a row appears to be a power of 2. Determination of the parity of a binomial coefficient was studied systematically by the British mathematician James Glaisher (1848-1928).

Theorem 4.1.10. *Let n and k be non-negative integers. Then*

$$\binom{n}{k} \equiv \begin{cases} 0 \bmod 2 & \text{if } n \text{ is even and } k \text{ is odd} \\ \binom{\lfloor n/2 \rfloor}{\lfloor k/2 \rfloor} \bmod 2 & \text{otherwise} \end{cases}$$

Proof: This proof splits naturally into four cases.

Case 1 — n even and k odd: Since n is even, it is clear that, for this case, the value of the right side of the absorption identity

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

is even. Since the product $k\binom{n}{k}$ on the left side must also be even, and since k is odd, it follows that $\binom{n}{k}$ is even.

Case 2 — n even and k even: For this case, we expand the binomial coefficient.

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{1\cdot 2\cdot 3\cdots k}$$

$$= \frac{(n-1)(n-3)\cdots(n-k+1)}{1\cdot 3\cdot 5\cdots(k-1)} \cdot \frac{n\,(n-2)(n-4)\cdots(n-k+2)}{2\cdot 4\cdot 6\cdots k}$$

Since the denominator has $k/2$ even factors, we continue

$$= \frac{(n-1)(n-3)\cdots(n-k+1)}{1\cdot 3\cdot 5\cdots(k-1)} \cdot \frac{n\,(n-2)(n-4)\cdots(n-k+2)}{2^{\frac{k}{2}}\cdot 1\cdot 2\cdot 3\cdots \frac{k}{2}}$$

and, since the numerator has $k/2$ even factors,

$$= \frac{(n-1)(n-3)\cdots(n-k+1)}{1\cdot 3\cdot 5\cdots(k-1)} \cdot \frac{2^{\frac{k}{2}}\cdot \frac{n}{2}(\frac{n}{2}-1)(\frac{n}{2}-2)\cdots(\frac{n}{2}-\frac{k}{2}+1)}{2^{\frac{k}{2}}\cdot 1\cdot 2\cdot 3\cdots \frac{k}{2}}$$

$$= \frac{(n-1)(n-3)\cdots(n-k+1)}{1\cdot 3\cdot 5\cdots(k-1)} \cdot \binom{n/2}{k/2}$$

Therefore,

$$1\cdot 3\cdot 5\cdots(k-1)\binom{n}{k} \;=\; (n-1)(n-3)\cdots(n-k+1)\binom{n/2}{k/2}$$

It follows that for n and k both even,

$$\binom{n}{k} \equiv \binom{n/2}{k/2} \equiv \binom{\lfloor n/2\rfloor}{\lfloor k/2\rfloor} \bmod 2 \qquad (4.1.11)$$

The first equivalence in (4.1.11) holds because each of the factors preceding the binomial coefficient in the numerator and in the denominator is odd, and multiplication of an integer by an odd number does not change its parity. The second holds because $n/2 = \lfloor n/2\rfloor$ and $k/2 = \lfloor k/2\rfloor$ for N and k both even.

Case 3 — n odd and k odd: As in Case 1, our starting point is the absorption identity

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

Since n and k are both odd, and once again, since multiplication of an integer by an odd number does not change the parity, it follows that

$$\binom{n}{k} \equiv \binom{n-1}{k-1} \bmod 2$$

Since $n-1$ and $k-1$ are both even, it follows from Case 2 that

$$\binom{n-1}{k-1} \equiv \binom{\lfloor n/2 \rfloor}{\lfloor k/2 \rfloor} \bmod 2$$

and, hence, that

$$\binom{n}{k} \equiv \binom{\lfloor n/2 \rfloor}{\lfloor k/2 \rfloor} \bmod 2$$

Case 4 — n odd and k even: The symmetry identity implies that

$$(n-k)\binom{n}{k} = (n-k)\binom{n}{n-k} \quad \text{and} \quad n\binom{n-1}{n-k-1} = n\binom{n-1}{k}$$

It follows from the absorption identity

$$(n-k)\binom{n}{n-k} = n\binom{n-1}{n-k-1}$$

that

$$(n-k)\binom{n}{k} = n\binom{n-1}{k}$$

Since $n-k$ and n are both odd, we have

$$\binom{n}{k} \equiv \binom{n-1}{k} \bmod 2$$

Applying Case 2 to the right side, we obtain

$$\binom{n}{k} \equiv \binom{\lfloor (n-1)/2 \rfloor}{\lfloor k/2 \rfloor} \bmod 2$$

Since n is odd, the upper index $\lfloor (n-1)/2 \rfloor$ equals $\lfloor n/2 \rfloor$. $\diamondsuit$

A simple algorithm to decide the parity of a binomial coefficient is to apply Theorem 4.1.10 iteratively, either until the upper index is even and the lower index odd or until the lower index is 0.

Example 4.1.5: Here are both possible types of termination.

$$\binom{165}{93} \equiv \binom{82}{46} \equiv \binom{41}{23} \equiv \binom{20}{11} \equiv 0 \bmod 2$$

$$\binom{75}{11} \equiv \binom{37}{5} \equiv \binom{18}{2} \equiv \binom{9}{1} \equiv \binom{4}{0} \equiv 1 \bmod 2$$

In order to see why the number of odd binary coefficients in a row of Pascal's triangle is a power of 2, we observe that, in binary numbers, the integer operation

$$n \mapsto \lfloor n/2 \rfloor$$

is achieved by erasing the rightmost bit. We observe also that Case 1 of Theorem 4.1.10, in which n is even and k is odd, is discernible by a 0-bit at the right end of the binary numeral for n and a 1-bit at the right end of the binary numeral for k. If the parity algorithm uses binary numerals, then iterative erasure of the rightmost bits is not actually necessary. It is possible, instead, to align both numerals flush right and to scan to see whether there is a 0-bit above a 1-bit.

Example 4.1.5, cont.: In scanning the aligned binary numerals

$$\begin{aligned} 165_{10} &= 1\,0\,1\,0\,0\,1\,0\,1_2 \\ 93_{10} &= 0\,1\,0\,1\,1\,1\,0\,1_2 \end{aligned}$$

leftward from the right end, the first occurrence of a 0 in the upper index occurs at the 2^1-bit. Since there is also a 0-bit immediately below it, the scan continues. The next 0 in the upper index occurs at the 2^3-bit, and there is a 1-bit below it, so the scan terminates and the decision is even parity. In scanning the aligned binary numerals

$$\begin{aligned} 75_{10} &= 1\,0\,0\,1\,0\,1\,1_2 \\ 11_{10} &= 0\,0\,0\,1\,0\,1\,1_2 \end{aligned}$$

one observes that there is a 0-bit beneath every 0-bit in the upper index, so the decision is odd parity.

Proposition 4.1.11. *The number of odd binomial coefficients in row n of Pascal's triangle is 2^w, where w is the number of 1-bits in the binary representation of n.*

Proof: For the binomial coefficient $\binom{n}{k}$ to be odd, there must be a 0 at each bit in the binary numeral for k for which there is a 0 at the corresponding bit of the binary numeral for n. However, if there is a 1 at a bit of the binary numeral for n, there may be either a 0 or a 1 at the corresponding bit of the binary numeral for k. If there are w 1-bits for n, then there are 2^w values for k that satisfy the rule for the 0-bits. $\diamondsuit$

Corollary 4.1.12. *If the integer n is of the form $2^r - 1$, then every binomial coefficient in row n of Pascal's triangle is odd.*

Proof: There are no 0-bits in the binary representation of $2^r - 1$. $\diamondsuit$

4.2 BINOMIAL INVERSION

This section develops an incremental technique used with binomial coefficients, called *binomial inversion.* Its main application in this section is within a solution of the derangement recurrence.

DEF: The ***transform*** of the sequence $\langle f_n \rangle$ under ***binomial inversion*** is the sequence $\langle g_n \rangle$ with

$$g_n = \sum_{j=0}^{n} \binom{n}{j} (-1)^j f_j \qquad (4.2.1)$$

A characteristic property of anything mathematical that is correctly called a ***duality operation*** is that a second application of the operation restores the original object. Theorem 4.2.1 confirms that a transformation called *binomial inversion* of sequences has this property.

Theorem 4.2.1. *Let $\langle f_n \rangle$ be a sequence and $\langle g_n \rangle$ its transform under binomial inversion. Then, for all $n \geq 0$,*

$$f_n = \sum_{j=0}^{n} \binom{n}{j} (-1)^j g_j \qquad (4.2.2)$$

In other words, retransformation restores the original sequence $\langle f_n \rangle$.

Proof: Start at the right side of Eq. (4.2.2) and substitute the inversion formula of Eq. (4.2.1) for g_j.

$$\sum_{j=0}^{n} \binom{n}{j} (-1)^j g_j = \sum_{j=0}^{n} \binom{n}{j} (-1)^j \sum_{i=0}^{j} \binom{j}{i} (-1)^i f_i$$

$$= \sum_{j=0}^{n} \sum_{i=0}^{j} \binom{n}{j} \binom{j}{i} (-1)^{j+i} f_i \qquad (4.2.3)$$

Exchanging the order of summation is useful here.

$$= \sum_{i=0}^{n} \sum_{j=i}^{n} \binom{n}{j} \binom{j}{i} (-1)^{j+i} f_i \qquad (4.2.4)$$

Applying subset-of-a-subset identity (Prop 4.1.8) reduces the # occurrences of the summation index j.

$$= \sum_{i=0}^{n} \sum_{j=i}^{n} \binom{n}{i} \binom{n-i}{j-i} (-1)^{j+i} f_i$$

Then factor to simplify the inner summation.

$$= \sum_{i=0}^{n} \binom{n}{i} f_i \sum_{j=i}^{n} \binom{n-i}{j-i} (-1)^{j-i} \text{ since } (-1)^{2i} = 1$$

Substitute $k = j - i$.

$$= \sum_{i=0}^{n} \binom{n}{i} f_i \sum_{k=0}^{n-i} \binom{n-i}{k} (-1)^k$$

Inner summation $\longrightarrow$ exponentiated binomial.

$$= \sum_{i=0}^{n} \binom{n}{i} f_i \, (1-x)^{n-i} \Big|_{x=1}$$

Using the Iverson truth function $(i = n)$ leads to completion of the proof.

$$= \sum_{i=0}^{n} \binom{n}{i} f_i \, (i = n)$$

$$= \binom{n}{n} f_n \, (n = n) \;=\; f_n \qquad \diamondsuit$$

Observe at Eq. (4.2.4) above that the summation index j occurs twice in the summand within a binomial coefficient, once as an upper index, and once as a lower index. In such circumstances, as seen here, the subset-of-a-subset identity often facilitates a transformation that reduces the number of occurrences of the summation index in the summand.

Some Basic Examples of Inversions

The first three examples here of inversion are introductory, to show how inversion works.

Example 4.2.1: The constant sequence

$$\langle f_n \rangle \;=\; 1 \quad 1 \quad 1 \quad 1 \quad \cdots$$

has the inversion

$$\begin{aligned} g_n &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j f_j \\ &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j \\ &= (1-x)^n \Big|_{x=1} \\ &= (1-1)^n \;=\; \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n > 0 \end{cases} \\ \Rightarrow \qquad \langle g_n \rangle &= 1 \quad 0 \quad 0 \quad 0 \quad \cdots \end{aligned}$$

More generally, the sequence

$$\langle f_n \rangle = c \quad c \quad c \quad c \quad \cdots$$

has the inversion

$$\langle g_n \rangle = c \quad 0 \quad 0 \quad 0 \quad \cdots$$

Example 4.2.2: The natural number sequence

$$\langle f_n \rangle = 0 \quad 1 \quad 2 \quad 3 \quad \cdots$$

is inverted as follows.

$$\begin{aligned} g_n &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j f_j \\ &= \sum_{j=0}^{n} j \binom{n}{j} (-1)^j \end{aligned} \tag{4.2.5}$$

Apply the absorption identity to eliminate an occurrence of the index j.

$$= \sum_{j=0}^{n} n \binom{n-1}{j-1} (-1)^j$$

$$= n \sum_{j=0}^{n} \binom{n-1}{j-1} (-1)^j$$

Substitute $j = i+1$ to align the binomial coefficient with the summation limits.

$$= n \sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^{i+1}$$

$$= -n \sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^{i}$$

$$= -n\,(1-x)^{n-1}\Big|_{x=1} = \begin{cases} -1 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}$$

$$\Rightarrow \quad \langle g_n \rangle = 0 \quad -1 \quad 0 \quad 0 \quad \cdots$$

In Eq. (4.2.5) of this example, the summation index j occurs within a binomial coefficient and also as a multiplier. The absorption identity is the usual binomial identity by which the number of occurrences of the index variable is reduced in such a circumstance.

The seq $0 \quad 1 \quad 2 \quad 3 \quad \cdots$ can also be represented as $\langle\binom{n}{1}\rangle$. Accordingly, it should be unsurprising if calculating the inversion of the sequence $\binom{n}{r}$ is similar to Example 4.2.2.

Example 4.2.3: The binomial coefficient sequence

$$f_n = \binom{n}{r}$$

for a fixed non-negative number r has the inversion sequence

$$\begin{aligned} g_n &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j f_j \\ &= \sum_{j=0}^{n} \binom{j}{r} \binom{n}{j} (-1)^j \qquad (4.2.6) \end{aligned}$$

Apply the subset-of-a-subset identity and then factor.

$$\begin{aligned} &= \sum_{j=0}^{n} \binom{n}{r} \binom{n-r}{j-r} (-1)^j \\ &= \binom{n}{r} \sum_{j=0}^{n} \binom{n-r}{j-r} (-1)^j \end{aligned}$$

Substitute $j = i + r$. Then

$$\begin{aligned} g_n &= \binom{n}{r} \sum_{i=-r}^{n-r} \binom{n-r}{i} (-1)^{i+r} \\ &= \binom{n}{r} \sum_{i=0}^{n-r} \binom{n-r}{i} (-1)^{i+r} \end{aligned}$$

$$= (-1)^r \binom{n}{r} \sum_{i=0}^{n-r} \binom{n-r}{i} (-1)^i$$

$$= \begin{cases} (-1)^n & \text{if } n = r \\ 0 & \text{if } n \neq r \end{cases}$$

At Eq. (4.2.6), the summand has two occurrences of the summation index j. This time, both are within different binomial coefficients, with one occurrence as an upper index and the other as a lower index. The subset-of-a-subset identity is frequently used to eliminate one of the occurrences in such summands, thereby simplifying the sum.

Derangements

The point of learning how to invert sequences is not just to pose a new class of computational exercises. Binomial inversion has numerous extrinsic applications.

Example 4.2.4: Every permutation of the integer interval $[1:n]$ can be obtained by choosing r numbers from $[1:n]$ and deranging them. Accordingly, if D_j is a derangement number, then

$$n! = \binom{n}{0} D_0 + \binom{n}{1} D_1 + \binom{n}{2} D_2 + \cdots + \binom{n}{n} D_n$$

It follows that the sequence

$$f_n = (-1)^n D_n$$

has the binomial inversion

$$g_n = n!$$

By the duality property of binomial inversion, we have

$$f_n = (-1)^n D_n = \sum_{j=0}^{n} \binom{n}{j} (-1)^j g_j$$

which implies that

$$\begin{aligned} D_n &= (-1)^n \sum_{j=0}^{n} \binom{n}{j} (-1)^j g_j \\ &= (-1)^n \sum_{j=0}^{n} \binom{n}{j} (-1)^j j! \\ &= \sum_{j=0}^{n} n^{\underline{j}} (-1)^j \\ \Rightarrow \frac{D_n}{n!} &= 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \underset{n\to\infty}{\xrightarrow{\lim}} e^{-1} \end{aligned}$$

Thus, the proportion of derangements among the permutations of a set of n objects tends to e^{-1} as n gets larger, a result that we previously derived with inclusion-exclusion in Example 3.6.5. This illustrates again our perspective that it is helpful to have a variety of mathematical tools available for the solution of a given problem.

More Examples of Inversions

The summation techniques presented in this section for transforming sequences are widely applicable. The next section of this chapter applies these methods to computations in probability and statistics. We complete the present section with two more examples that combine the method of binomial inversion with the binomial identities derived previously.

Example 4.2.5: When two factors of a summand are both binomial coefficients that contain the index of summation as a lower index, the key to simplification is to set up an application of the Vandermonde convolution, which would simplify the summand. The sequence

$$f_n = (-1)^n \binom{N}{n}$$

has as its binomial inversion the sequence

$$\begin{aligned} g_n &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j f_j \\ &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j (-1)^j \binom{N}{j} \\ &= \sum_{j=0}^{n} \binom{N}{j} \binom{n}{j} \end{aligned}$$

Apply the symmetry identity as a setup

$$= \sum_{j=0}^{n} \binom{N}{j} \binom{n}{n-j}$$

and then invoke the Vandermonde convolution.

$$= \binom{N+n}{n}$$

Example 4.2.6: Sometimes there is a quotient of two binomial coefficients both of which contain the index of summation. The sequence

$$f_n = (-1)^n \binom{N}{n}^{-1}$$

has as its transform under binomial inversion the sequence

$$\begin{aligned} g_n &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j f_j \\ &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j (-1)^j \binom{N}{j}^{-1} \\ &= \sum_{j=0}^{n} \frac{\binom{n}{j}}{\binom{N}{j}} \end{aligned}$$

Here we apply the subset-of-a-subset identity

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

thereby obtaining

$$
\begin{aligned}
g_n &= \sum_{j=0}^{n} \frac{\binom{n}{j}}{\binom{N}{n}\binom{n}{j} \Big/ \binom{N-j}{n-j}} \\
&= \binom{N}{n}^{-1} \sum_{j=0}^{n} \binom{N-j}{n-j}
\end{aligned}
$$

which can be simplified using the diagonal-sum identity

$$
\begin{aligned}
&= \binom{N}{n}^{-1} \binom{N+1}{n} \\
&= \frac{N+1}{N-n+1}
\end{aligned}
$$

4.3 APPLICATIONS TO STATISTICS

Binomial coefficients frequently occur in sums that arise in probability and statistics. We continue to seek to reduce the number of occurrences of the index of summation within the summand. There are a few additional rules of thumb to be learned here and used.

Probability and Random Variables

Some basic definitions are now recalled from elementary statistics and probability. The pace of the exposition here presumes that the reader has some prior familiarity with these topics.

DEF: A ***discrete probability space*** is a pair $\langle \Omega, \Pr \rangle$ as follows.

- The discrete set Ω is called a ***sample space***.
- A subset of Ω is called an ***event***.
- The set 2^Ω of all subsets of Ω is called the ***event space***.
- The function $\Pr : 2^\Omega \to \mathbb{R}$ is called a ***probability measure***, and it satisfies the following axioms.
    1. $0 \leq \Pr(A) \leq 1$, for every event $A \subseteq \Omega$. The number $\Pr(A)$ is called the ***probability of the event*** A.
    2. $\Pr(\Omega) = 1$.

3. If the events A_s, for $s \in S$, are mutually exclusive subsets of Ω, then

$$\Pr\Big(\bigcup_{s\in S} A_s\Big) = \sum_{s\in S} \Pr(A_s)$$

DEF: A ***random variable*** X on a sample space is a real-valued function. It is called a ***discrete random variable*** if the set of values it takes is finite or countably infinite.

NOTATION: Let $X : \Omega \to \mathbb{R}$ be a discrete random variable on a sample space Ω with probability measure Pr. For $x \in \mathbb{R}$, the probability of the set $\{\omega \in \Omega \mid X(\omega) = x\}$ is denoted $\Pr(x)$.

Mean and Variance

The *expected value* of a random variable, also called the *mean*, is commonly described as a weighted average. The *variance* and the *standard deviation* are measures of dispersion from the mean.

DEF: Let $X : \Omega \to \mathbb{R}$ be a discrete random variable on a sample space Ω with probability measure Pr, and let D be the set of values that X takes. The ***expected value*** or ***mean of the random variable*** X, denoted $E(X)$ or μ_X, is the sum

$$E(X) = \mu_X = \sum_{x\in D} x \cdot \Pr(x) \qquad (4.3.1)$$

DEF: Let $X : \Omega \to \mathbb{R}$ be a discrete random variable on a sample space Ω with probability measure Pr, and let D be the set of values that X takes. The ***variance of the random variable*** X, denoted $V(X)$ or σ_X^2, is the sum

$$V(X) \;=\; \sigma_X^2 \;=\; \sum_{x \in D} (x - \mu_X)^2 \cdot \Pr(x) \;=\; E([X - \mu_X]^2) \tag{4.3.2}$$

DEF: Let $X : \Omega \to \mathbb{R}$ be a discrete random variable. The ***standard deviation of the random variable*** X, denoted $SD(X)$ or σ_X, is the square root of the variance.

$$SD(X) \;=\; \sigma_X \;=\; \sqrt{\sigma_X^2} \tag{4.3.3}$$

NOTATION: When it is clear from context to which random variable X they pertain, the subscripts for mean and variance may be denoted μ and σ^2.

DEF: In calculating the ***mean of a list of numbers*** or the ***variance of a list of numbers***, one regards each element of the list as equally likely.

Proposition 4.3.1. *Let $X : \Omega \to \mathbb{R}$ be a discrete random variable. Then*

$$\sigma_X^2 \;=\; E(X^2) - \mu^2 \tag{4.3.4}$$

Proof: Let D be the set of values taken by X. We proceed straightforwardly, starting from the Equation (4.3.2).

$$\sigma_X^2 \;=\; \sum_{x \in D} (x - \mu_X)^2 \cdot \Pr(x)$$

$$
\begin{aligned}
&= \sum_{x \in D} (x^2 - 2x\mu_X + \mu_X^2) \Pr(x) \\
&= \sum_{x \in D} x^2 \Pr(x) - \sum_{x \in D} 2x\mu_X \Pr(x) + \sum_{x \in D} \mu_X^2 \Pr(x) \\
&= E(X^2) - 2\mu_X \cdot \sum_{x \in D} x \Pr(x) + \mu_X^2 \cdot \sum_{x \in D} \Pr(x) \\
&= E(X^2) - 2\mu_X \cdot \mu_X + \mu_X^2 \cdot 1 \\
&= E(X^2) - \mu_X^2 \qquad \diamondsuit
\end{aligned}
$$

Binomial Distribution

The prototypical experiment whose outcomes have a binomial distribution is a sequence of n tosses of a coin. Taking one of the possible outcomes of an individual toss, say heads, to be "success", what is binomially distributed is the number of heads. We now apply the binomial coefficient identities of §4.1 to the calculation of the mean and variance of the binomial distribution.

DEF: Given an experiment with binary outcome (*success* or *failure*) that is to be performed n times, the ***binomial random variable*** X is the number of successes. Suppose that the probability of success is p, and that the n trials are independent. Then

$$Pr(X = j) \;=\; \binom{n}{j} p^j \, (1-p)^{n-j} \tag{4.3.5}$$

The sample space is the sequence of outcomes of the n trials. An event is a set of possible outcomes.

Proposition 4.3.2. *The expected value of a binomial random variable X on n trials, each with probability p of success, is*

$$E(X) = np$$

Proof: Eq. (4.3.1) defines expected value.

$$E(X) = \sum_{j=0}^{n} j \cdot \Pr(X = j)$$

We substitute the probability of a binomial random variable, as given by Eq. (4.3.5).

$$= \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j}$$

Absorption eliminates one of the four occurrences of the summation index j.

$$= \sum_{j=0}^{n} n \binom{n-1}{j-1} p^j (1-p)^{n-j}$$

$$= np \sum_{j=0}^{n} \binom{n-1}{j-1} p^{j-1} (1-p)^{n-j}$$

Substituting $i = j - 1$ yields the summation

$$= np \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i}$$

that we recognize as a binomial expansion, and simplify.

$$
\begin{aligned}
&= np\,[p + (1-p)]^{n-1} \\
&= np
\end{aligned}
\qquad \diamondsuit
$$

Prop 4.3.3. *The variance of a binomial random variable X on n trials, each with probability p of success, is*

$$V(X) \;=\; np(1-p)$$

Proof: Once again, start at Eq. (4.3.1).

$$
\begin{aligned}
E(X^2) &= \sum_{j=0}^{n} j^2 \cdot \Pr(X=j) \\
&= \sum_{j=0}^{n} j^2 \binom{n}{j} p^j \, (1-p)^{n-j}
\end{aligned}
$$

There are once again four occurrences of the index j of summation. Applying absorption reduces the exponent of j in one occurrence, a reasonable step.

$$
\begin{aligned}
&= \sum_{j=0}^{n} jn \binom{n-1}{j-1} p^j (1-p)^{n-j} \\
&= np \sum_{j=0}^{n} j \binom{n-1}{j-1} p^{j-1} (1-p)^{n-j}
\end{aligned}
$$

Substitute $i = j-1$ to align the indices of the binomial coefficient with the upper and lower limits of the summation, another reasonable step.

$$= np\sum_{i=0}^{n-1}(1+i)\binom{n-1}{i}p^i(1-p)^{n-1-i}$$

Splitting the sum like this helps here

$$= np\sum_{i=0}^{n-1}\binom{n-1}{i}p^i(1-p)^{n-1-i} \;+\; np\sum_{i=0}^{n-1} i\binom{n-1}{i}p^i(1-p)^{n-1-i}$$

because the summation in the first part is recognizable as a binomial expansion.

$$= np \;+\; np\sum_{i=0}^{n-1} i\binom{n-1}{i}p^i(1-p)^{n-1-i}$$

Applying absorption again now eliminates one occurrence of the summation index.

$$= np \;+\; np\sum_{i=0}^{n}(n-1)\binom{n-2}{i-1}p^i(1-p)^{n-1-i}$$

Substituting $k = i-1$ realigns the lower index of the binomial coefficient with the lower limit of the summation.

$$= np \;+\; n(n-1)\,p^2\sum_{k=0}^{n}\binom{n-2}{k}p^k(1-p)^{n-2-k}$$

$$= np \;+\; n(n-1)\,p^2 \;=\; np + n^2p^2 \;-\; np^2$$

$$= n^2p^2 \;+\; np(1-p)$$

By Propositions 4.3.1 and 4.3.2,

$$\sigma_X^2 \;=\; E(X^2) \;-\; E(X)^2$$

$$\begin{aligned} &= [n^2p^2 + np(1-p)] - n^2p^2 \\ &= np(1-p) \end{aligned} \qquad \diamondsuit$$

Unbiased Estimator of the Mean

An intuitive statistical approach to estimating the proportion of persons in a population of large size N who have a given characteristic (such as enjoying recreational mathematics) is to take a random sample and to use the proportion in that sample to estimate the proportion in the general population. We will use binomial coefficient identities in confirming the validity of this approach.

DEF: An estimator $\hat{\theta}$ of a statistical characteristic θ of a population is said to be an ***unbiased estimator*** if the expected value $E(\hat{\theta})$ for a random sample equals θ.

Proposition 4.3.4. *The sample proportion is an unbiased estimator of the proportion of individuals in a general population that have a given characteristic.*

Proof: Suppose that in a population of size N exactly M individuals have a given trait. A sample of size n is taken. The random variables of interest are the number m of persons with that trait and the proportion

$$X = \frac{m}{n}$$

of persons with the trait. The total number of ways to choose a sample of size n is

$$\binom{N}{n}$$

The number of ways that a sample of size n could have exactly j persons with the prescribed trait is the product

$$\binom{M}{j}\binom{N-M}{n-j}$$

of the number of choices of j individuals from the population of size M with the trait and the number of choices of the remaining $n-j$ individuals from the $N-M$ persons who do not have the trait. Thus,

$$Pr\,(m=j) \;=\; \frac{\binom{M}{j}\binom{N-M}{n-j}}{\binom{N}{n}}$$

Accordingly,

$$\begin{aligned} E(X) &= \sum_{j=0}^{n} \frac{j}{n}\cdot Pr\,(m=j) \\ &= \frac{1}{n}\sum_{j=0}^{n} j\cdot \frac{\binom{M}{j}\binom{N-M}{n-j}}{\binom{N}{n}} \end{aligned}$$

$$= \frac{1}{n}\binom{N}{n}^{-1}\sum_{j=0}^{n} j\cdot\binom{M}{j}\binom{N-M}{n-j}$$

Apply the absorption identity to eliminate one occurrence of the index of summation.

$$= \frac{1}{n}\binom{N}{n}^{-1}\sum_{j=0}^{n} M\cdot\binom{M-1}{j-1}\binom{N-M}{n-j}$$

$$= \frac{M}{n}\binom{N}{n}^{-1}\sum_{j=0}^{n}\binom{M-1}{j-1}\binom{N-M}{n-j}$$

Now use the Vandermonde convolution.

$$= \frac{M}{n}\binom{N}{n}^{-1}\binom{N-1}{n-1}$$

$$= \frac{M}{n}\cdot\frac{n!}{N^{\underline{n}}}\cdot\frac{(N-1)^{\underline{n-1}}}{(n-1)!} \;=\; \frac{M}{N}$$

Thus, the intuitive method of estimating the mean is unbiased. $\diamondsuit$

Unbiased Estimator of the Variance

Let X be a random variable on a space Ω. The identically distributed random variables

$$X_1 \quad X_2 \quad \ldots \quad X_n$$

are the values of X on n samples from Ω, with sample mean $\overline{X}$. Statisticians use the estimator

$$\hat{\sigma^2} = \frac{\sum(X_i - \overline{X})^2}{n-1} = \frac{\sum X_i^2 - n^{-1}\left(\sum X_i\right)^2}{n-1}$$

with $n-1$ in the denominator (rather than n), for the variance. This is explained by the next proposition.

Proposition 4.3.5. *The sample statistic*

$$\hat{\sigma^2} = \frac{\sum(X_i - \overline{X})^2}{n-1} = \frac{\sum X_i^2 - n^{-1}\left(\sum X_i\right)^2}{n-1} \tag{4.3.6}$$

is an unbiased estimator of the variance of the random variable X.

Proof:

$$\begin{aligned} E\left(\hat{\sigma^2}\right) &= \frac{E\left(\sum X_i^2\right)}{n-1} - \frac{E\left[\left(\sum X_i\right)^2\right]}{n(n-1)} \\ &= \frac{1}{n-1}\sum_{i=1}^{n} E(X_i^2) - \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i X_j) \end{aligned}$$

Split the double summation into two parts.

$$\begin{aligned} &= \frac{1}{n-1}\sum_{i=1}^{n} E(X_i^2) - \frac{1}{n(n-1)}\sum_{i=1}^{n} E(X_i^2) \\ &\qquad - \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i X_j)\cdot(j \neq i) \end{aligned}$$

$$= \left(\frac{1}{n-1} - \frac{1}{n(n-1)}\right) \cdot \sum_{i=1}^{n} E(X^2)$$

$$- \frac{1}{n(n-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} E(X)E(X) \cdot (j \neq i)$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} E(X^2) - \frac{1}{n(n-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} E(X)E(X) \cdot (j \neq i)$$

Both sums are resolvable.

$$= \frac{1}{n} \cdot nE(X^2) - \frac{n(n-1)}{n(n-1)} E(X)^2$$

$$= E(X^2) - E(X)^2$$

$$= V(X)$$

Thus, division by $n-1$ leads to an unbiased estimate. $\diamondsuit$

4.4 THE CATALAN RECURRENCE

REVIEW FROM §1.2:

- The ***Catalan sequence*** $\{c_n\}$ is defined by the recurrence

$$\begin{aligned} c_0 &= 1; && \text{initial value} \\ c_n &= c_0c_{n-1} + c_1c_{n-2} + \cdots + c_{n-1}c_0 && \text{for } n \geq 1 \end{aligned}$$

Surprisingly, perhaps, the closed formula for c_n is a multiple of a binomial coefficient.

Binary Trees

In graph theory, the set $\mathcal{T}_{\downarrow}$ of ***binary trees*** can be defined recursively:

- The empty tree Φ is in the set $\mathcal{T}_{\downarrow}$.
- The tree $K_1^{\bullet}$ with a single vertex designated as the *root* is in the set $\mathcal{T}_{\downarrow}$.
- If $T \in \mathcal{T}_{\downarrow}$, and if v is a vertex of the tree T, then each of the following rooted trees is in the set $\mathcal{T}_{\downarrow}$.
  i. The tree obtained by adjoining a new vertex to v, called the *left-child* of vertex v. (A vertex has at most one left-child.)
  ii. The tree obtained by adjoining a new vertex to v, called the *right-child* of vertex v. (A vertex has at most one right-child.)

Figure 4.4.1 illustrates the binary trees with 0, 1, 2, and 3 vertices. It is easy enough to verify for these small cases that the Catalan number c_n is the number of binary trees with n vertices.

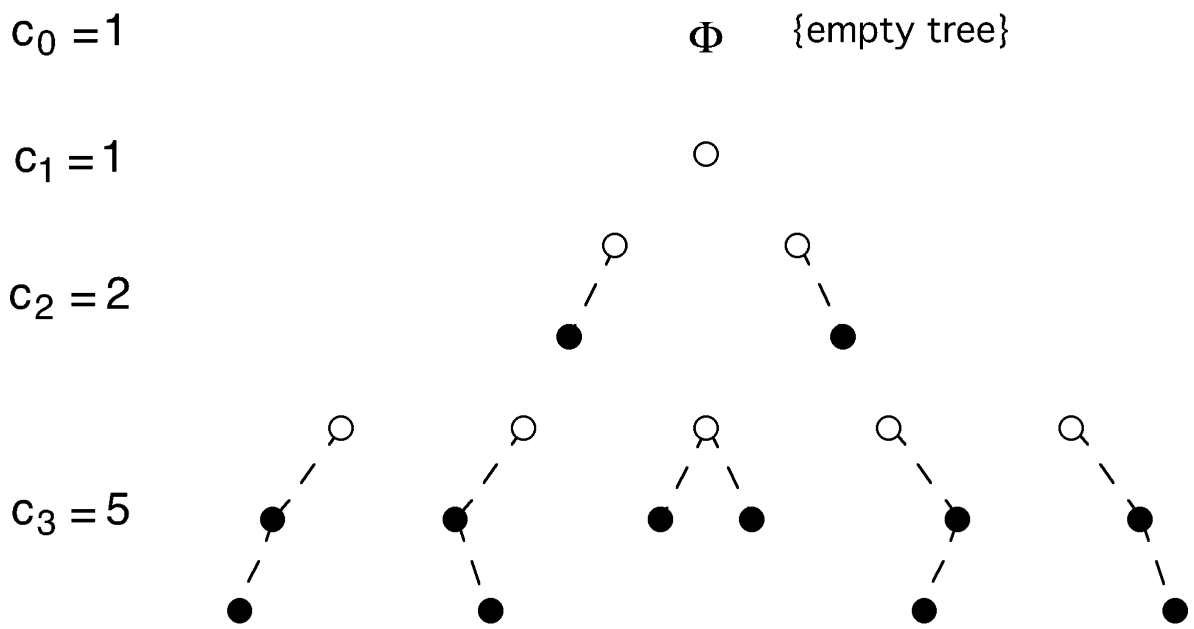

Fig 4.4.1 The smallest binary trees.

Remark: In computer science, each child of a vertex of a binary tree is designated either as a left-child or a right-child, even if there is only one child. The importance of this designation occurs in applications such as *binary search trees* and *priority trees* (see [GrYe2006]).

DEF: The ***left subtree*** of a binary tree T is the subtree whose root is the left-child of the root of T. The ***right subtree*** of a binary tree T is the subtree whose root is the right-child of the root of T.

Proposition 4.4.1. *For $n \geq 0$, the number of n-vertex binary trees equals the Catalan number c_n.*

Proof: By induction on the number n of vertices.

BASIS: Clearly, $c_0 = 1$ and $c_1 = 1$ are the numbers of binary trees with 0 and 1 vertices, respectively.

IND HYP: Let $n > 0$. Suppose for all integers k with $0 \leq k < n$, that c_k is the number of binary trees with k vertices.

IND STEP: Suppose that a binary tree has n vertices. For $k = 0, 1, \ldots, n-1$, the number of possible left subtrees with k vertices is c_k, by the induction hypothesis. Of course, the right subtree would then have $n-k-1$ vertices, so that there would be a total number of n vertices within the union of the two subtrees and the root, as depicted in Figure 4.4.2.

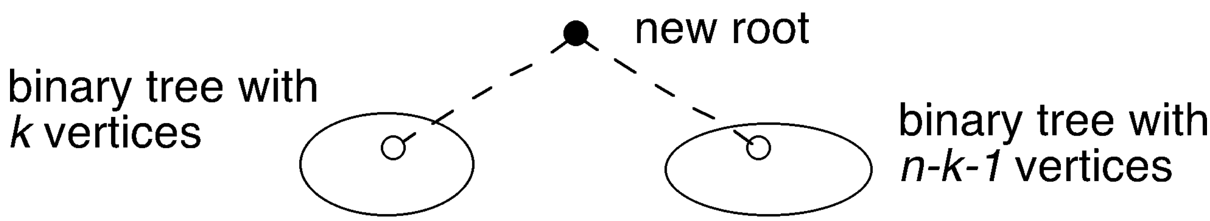

Fig 4.4.2 Joining left and right subtrees to a root.

The induction hypothesis also implies, therefore, that the number of possible right subtrees is c_{n-k-1}. Hence, by the rule of product, there are $c_k c_{n-k-1}$ n-vertex binary trees with k vertices in the left subtree. Accordingly, the total number of n-vertex binary trees is given by the sum

$$\begin{aligned} c_n &= \sum_{k=0}^{n-1} c_k\, c_{n-k-1} \\ &= c_0\, c_{n-1} + c_1\, c_{n-2} + \cdots + c_{n-1}\, c_0 \end{aligned} \qquad \diamondsuit$$

Nested Parentheses

The set $\mathcal{P}$ of ***well-nested strings of parentheses*** is defined recursively (as depicted in Figure 4.4.3 below):

- The empty string Λ is in $\mathcal{P}$.
- If $P_i, P_o \in \mathcal{P}$, then the string $(P_i)P_o$ is in $\mathcal{P}$. That is, we insert the string P_i inside a new pair and then juxtapose the string P_o at the right.

In listing the well-nested strings with 0, 1, 2, and 3 pairs of parentheses, the new pair specified by the recursion rule above is depicted by brackets.

0 pairs	Λ = empty string	$c_0 = 1$
1 pair	[ ]	$c_1 = 1$
2 pairs	[ ](), [()]	$c_2 = 2$
3 pairs	[ ]()(), [ ](()),[()](), [()()], [(())]	$c_3 = 5$

Proposition 4.4.2. *For $n \geq 0$, the number of well-nested strings of parentheses equals the Catalan number c_n.*

Proof: This proof follows the exact same lines as the proof of Proposition 4.4.1. The new pair of parentheses with well-nested substrings inside and outside in the recursive construction here corresponds to the new root with left and right binary subtrees there. ♢

Subdiagonal Paths

DEF: A ***northeastward path*** or ***NE-path*** in the array $[0:n] \times [0:n]$ is a path whose directed edges are each one unit in length and lead northward or eastward.

DEF: A ***subdiagonal path*** from $(0,0)$ to (n,n) in $[0:n] \times [0:n]$ is a NE-path along which each point (x,y) satisfies the inequality $x \geq y$.

The inequality in the definition means, as illustrated in Figure 4.4.3, that the path never crosses above the longest northeast diagonal.

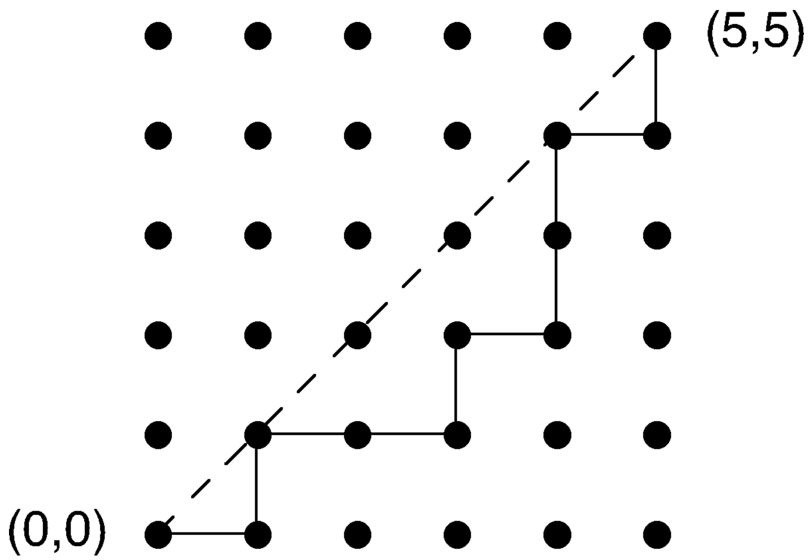

Fig 4.4.3 **A subdiagonal path from $(0,0)$ to (n,n).**

Proposition 4.4.3. *For $n \geq 0$, the number of subdiagonal paths from $(0,0)$ to (n,n) in the array $[0:n] \times [0:n]$ equals the Catalan number c_n.*

Proof: In every prefix of a well-nested string of parentheses, the number of left parentheses is greater than or equal to the number of right parentheses, and the total number of left parentheses equals the total number of right parentheses. Both these facts are provable by an induction on the length of the string. It follows that the well-nested strings of n pairs of parentheses are in bijective correspondence with the subdiagonal paths in the array $[0:n] \times [0:n]$. ♢

Solving the Catalan Recurrence

Of the many methods of solving the Catalan recurrence, the one now presented, based on work of D. Andre in 1878, is probably the simplest to follow.

Theorem 4.4.4. *The Catalan recurrence*

$$\begin{aligned} c_0 &= 1; && \text{initial value} \\ c_n &= c_0 c_{n-1} + c_1 c_{n-2} + \cdots + c_{n-1} c_0 && \textit{for } n \geq 1 \end{aligned}$$

has the solution

$$c_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{n+1} \cdot \frac{(2n)!}{n!\,n!}$$

Proof: Consider the set S_{NE} of all NE paths from $(0,0$ to (n,n) in $[0:n] \times [0:n]$. Suppose that each step eastward on an NE path is represented by the letter E and that each step northward is represented by the letter N. For instance, the path in Figure 4.4.3 is represented by the string

$$\text{ENEENENNEN}$$

This correspondence is evidently a bijection between the set S_{NE} of NE-paths and the set of strings in E and N of length $2n$ with n occurrences of each letter. The number of ways to choose the n locations for the N's in such a string is

$$\binom{2n}{n}$$

The bijection establishes that this is the total number of NE paths. By Proposition 4.4.3, the Catalan number c_n equals the number of subdiagonal paths in the array $[0:n] \times [0:n]$. Our approach is to subtract from this total the number of strings that do *not* represent subdiagonal paths.

Observe that a path is not subdiagonal if and only if, at some point, the number of northward steps has exceeded the number of eastward steps. Accordingly, the corresponding string

$$s_1 s_2 \ldots s_{2n}$$

would have a prefix in which the number of N's exceeds the number of E's. If $2j+1$ is the smallest index at which this occurs, then the number of E's in the prefix

$$s_1 s_2 \ldots s_{2j+1}$$

is j and the number of N's is $j+1$.

It follows that in the suffix

$$s_{2j+2}s_{2j+3}\cdots s_{2n}$$

there are $n-j$ E's and $n-j-1$ N's. Suppose that in the suffix, each E is replaced by an N and each N by an E. This is called a ***reflection of the subpath*** or a ***reflection of the substring***. The resulting string has $n-1$ E's and $n+1$ N's. It represents an NE path in $[0:n-1]\times[0:n+1]$, and there is a bijection between the set of non-subdiagonal paths from $(0,0)$ to (n,n) in $[0:n]\times[0:n]$ and the set of NE paths from$(0,0)$ to $(n-1,n+1)$ in $[0:n-1]\times[0:n+1]$, whose cardinality is

$$\binom{2n}{n-1}$$

Thus, the number of subdiagonal paths from $(0,0)$ to (n,n) in $[0:n]\times[0:n]$ is

$$\begin{aligned}\binom{2n}{n}-\binom{2n}{n-1} &= \frac{(2n)^{\underline{n-1}}(n+1)}{n!}-\frac{(2n)^{\underline{n-1}}n}{n!}\\ &= \frac{1}{n+1}\binom{2n}{n} \qquad \diamondsuit\end{aligned}$$

Example 4.4.1: $c_3 = \frac{1}{4}\binom{6}{3} = \frac{20}{4} = 5.$

Example 4.4.2: $c_4 = \frac{1}{5}\binom{8}{4} = \frac{70}{5} = 14.$

Generalized Binomial Theorem

An alternative proof of the solution to the Catalan recurrence uses the *generalized binomial theorem.*

NOTATION: The k^{th} derivative of a function $f(x)$ may be denoted $f^{(k)}$.

DEF: An *analytic function* is a function $f(x)$ with an n^{th} derivative for every $n \geq 0$.

Thm 4.4.5 (Generalized Binomial Theorem). *For any real number s, the exponentiated binomial $(1+x)^s$ has the power series*

$$\begin{aligned} (1+x)^s &= \sum_{k=0}^{\infty} \binom{s}{k} x^k \\ &= 1 + \frac{s^{\underline{1}}}{1!} \cdot x + \frac{s^{\underline{2}}}{2!} \cdot x^2 + \cdots + \frac{s^{\underline{k}}}{k!} \cdot x^k + \cdots \end{aligned}$$

Proof: For $f(x) = (1+x)^s$, observe that

$$\begin{aligned} f(0) &= (1+x)^s \Big|_{x=0} = 1^s = s^{\underline{0}} \\ f'(0) &= s(1+x)^{s-1} \Big|_{x=0} = s \cdot 1^{s-1} = s^{\underline{1}} \\ f''(0) &= s^{\underline{2}}(1+x)^{s-2} \Big|_{x=0} = s^{\underline{2}} \cdot 1^{s-2} = s^{\underline{2}} \end{aligned}$$

By induction, it can be proved that

$$f^{(k)}(0) = s^{\underline{k}}$$

Recall that the Maclaurin series expansion* of an analytic function $f(x)$ is

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \cdot x^k \\ &= \frac{f(0)}{0!} \cdot x^0 \; + \; \frac{f^{'}(0)}{1!} \cdot x^1 \; + \; \frac{f^{''}(0)}{2!} \cdot x^2 \; + \; \cdots \end{aligned}$$

Thus, the substitution $f^{(n)}(0) = s^{\underline{n}}$ yields the conclusion. $\diamondsuit$

Example 4.4.3: In the solution of the Catalan recurrence below, we use this generalized binomial expansion.

$$\begin{aligned} (1-4z)^{1/2} &= \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} (-4z)^k \\ &= 1 \; + \; \frac{(\frac{1}{2})^{\underline{1}}}{1!} \cdot (-4z) \; + \; \frac{(\frac{1}{2})^{\underline{2}}}{2!} \cdot (-4z)^2 \; + \; \cdots \end{aligned}$$

---

* This is equivalent to the Taylor series expansion at $x = 0$.

Alternative Proof of the Catalan Formula

An alternative proof of the solution

$$c_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{n+1}\cdot\frac{(2n)!}{n!n!}$$

to the Catalan recurrence provides a traditional illustration of the power of the method of generating functions in solving recurrences. We define the generating function

$$C(z) = \sum_{n=0}^{\infty} c_n z^n$$

and begin as in §2.2.

Step 1a. Multiply both sides of the Catalan recursion by z^n.

$$c_n z^n = \sum_{k=0}^{n-1} c_k c_{n-k-1} z^n \tag{4.4.1}$$

Step 1b. Sum both sides of Eq. (4.4.1) over the same range of values, as large as possible.

$$\sum_{n=1}^{\infty} c_n z^n = \sum_{n=1}^{\infty}\sum_{k=0}^{n-1} c_k c_{n-k-1} z^n \tag{4.4.2}$$

Step 2. Replace the infinite sum on the left of Eq. (4.4.2) with a finite sum involving the generating function $C(z)$.

$$C(z) - c_0 = \sum_{n=1}^{\infty}\sum_{k=0}^{n-1} c_k c_{n-k-1} z^n$$

Exchange the order of summation.

$$= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} c_k c_{n-k-1} z^n$$

$$= z \sum_{k=0}^{\infty} c_k z^k \sum_{n=k+1}^{\infty} c_{n-k-1} z^{n-k-1}$$

Substitute $j = n - k - 1$.

$$= z \sum_{k=0}^{\infty} c_k z^k \sum_{j=0}^{\infty} c_j z^j$$

$$= z \sum_{k=0}^{\infty} c_k z^k \, C(z)$$

$$= z \, C(z) \sum_{k=0}^{\infty} c_k z^k$$

$$= z \, C(z)^2$$

$$\Rightarrow \qquad zC(z)^2 - C(z) + 1 = 0 \qquad (4.4.3)$$

Step 3. Solve for $C(z)$ in Eq. (4.4.3) by the quadratic formula.

$$C(z) = \frac{1 \pm \sqrt{1-4z}}{2z} \qquad (4.4.4)$$

Step 4. To solve for the value of the general Catalan number c_n, we apply the Generalized Binomial Theorem, as in Example 4.4.3, to Eq. (4.4.4).

$$(1-4z)^{1/2} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-4z)^n = 1 + \sum_{n=1}^{\infty} \frac{\left(\frac{1}{2}\right)^{\underline{n}}}{n!} (-4z)^n$$

$$= 1 + \frac{(\frac{1}{2})^{\underline{1}}}{1!} \cdot (-4z) + \frac{(\frac{1}{2})^{\underline{2}}}{2!} \cdot (-4z)^2 + \cdots$$

Since every term of this series except the first is signed negative, the appropriate choice is the negative root. That is,

$$C(z) = \frac{1 - \sqrt{1 - 4z}}{2z} = \frac{-1}{2z} \sum_{n=1}^{\infty} \frac{(\frac{1}{2})^{\underline{n}}}{n!} (-4z)^n \qquad (4.4.5)$$

To simplify (4.4.5), we expand part of the summand

$$\begin{aligned}
\frac{1}{n!}\left(\frac{1}{2}\right)^{\underline{n}} &= \frac{1}{n!} \cdot \frac{1}{2} \cdot \frac{-1}{2} \cdot \frac{-3}{2} \cdot \cdots \cdot \frac{-(2n-3)}{2} \\
&= \frac{1}{n!} \cdot \frac{(-1)^{n-1}}{2^n} \prod_{j=1}^{n-1} (2j-1) \\
&= \frac{1}{n!} \cdot \frac{(-1)^{n-1}}{2^n} \prod_{j=1}^{n-1} \frac{(2j-1)(2j)}{2j} \\
&= \frac{1}{n!} \cdot \frac{(-1)^{n-1}}{2^n} \cdot \frac{(2n-2)!}{2^{n-1}(n-1)!} \\
&= \frac{(-1)^{n-1}(2n-2)!}{2^{2n-1}(n-1)!n!} \qquad (4.4.6)
\end{aligned}$$

and we substitute the result (4.4.6) back into Eq. (4.4.5), to obtain

$$C(z) = \frac{-1}{2z} \sum_{n=1}^{\infty} \frac{(\frac{1}{2})^{\underline{n}}}{n!} (-4z)^n$$

$$
\begin{aligned}
&= \frac{-1}{2z}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}(2n-2)!}{2^{2n-1}(n-1)!n!}(-4z)^n \\
&= \frac{-1}{2z}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}(2n-2)!}{2^{2n-1}(n-1)!n!}(-4z)^n \\
&= \frac{-1}{2z}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}(2n-2)!}{2^{2n-1}(n-1)!n!}(-1)^n 2^{2n} z^n \\
&= \sum_{n=1}^{\infty}\frac{(2n-2)!}{(n-1)!n!}z^{n-1} \;=\; \sum_{n=1}^{\infty}\frac{1}{n}\binom{2n-2}{n-1}z^{n-1} \\
&= \sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}z^n
\end{aligned}
$$

This yields the conclusion

$$c_n \;=\; \frac{1}{n+1}\binom{2n}{n} \qquad \diamondsuit$$