Binomial Coefficients

We will look at:

- simple properties and transformations
- some summations (that contain binomial coefficients)
- approximations

Binomial Theorem

\[(1+x)^n = \sum_{0 \leq i \leq n} \binom{n}{i} x^i \quad n - \text{natural number}\]

Another form (that appears more general) is:

\[(a+b)^n = \sum_{0 \leq i \leq n} \binom{n}{i} a^i b^{n-i}\]

In fact the two formulas are exactly the same, because we can write

\[b^n \left(1 + \frac{a}{b}\right)^n = b^n \sum_{0 \leq i \leq n} \binom{n}{i} \left(\frac{a}{b}\right)^i = \sum_{0 \leq i \leq n} \binom{n}{i} a^i b^{n-i}\]

The binomial coefficient \(\binom{n}{i}\) can be rendered in different ways:

\[
\binom{n}{i} = \begin{cases} 
\frac{n!}{i!(n-i)!} & 0 \leq i \leq n, \ n \in \mathbb{N} \ N = \{0, 1, 2 \ldots \} \\
\frac{n!}{i!} & n \text{ is any number, } n^k \text{ is “falling factorial” } = (n-i+1)(n-i+2)\ldots(n-1)n \\
0 & i \notin \mathbb{N}, i < 0, i > n \in \mathbb{N}
\end{cases}
\]

Another look at the falling factorial: \(n^i = \frac{n!}{(n-i)!} = \frac{1 \cdot 2 \cdot 3 \ldots n}{1 \cdot 2 \cdot \ldots \cdot (n-i)}\).

Transformations

- For \(n\) and \(i\) that are natural numbers,

\[\binom{n}{i} = \binom{n}{n-i}\]

Proof:

\[\binom{n}{i} = \frac{n!}{(n-i)!i!}, \quad \binom{n}{n-i} = \frac{n!}{(n-(n-i))!(n-i)!} = \frac{n!}{i!(n-i)!} \quad i, n \in \mathbb{N}\]
note: \( (1/3)^2 = \frac{1}{9} = 1/9 \).

\[
\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1} \quad n \in \mathbb{R}, k \in \mathbb{N}
\]

proof:

\[
\binom{n}{k} = \frac{n^k}{k!} = \frac{n(n-1)(n-k+1)\ldots(n-k+1)}{k(k-1)\ldots1} = \frac{n(n-k)^{k-1}}{k(k-1)!}
\]

**Defining recursion**

\[
\binom{n}{0} = \frac{n}{0} = 1 \quad n \geq 0
\]

\[
\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1} \quad 0 < i < n
\]

proof: \((1+x)^n = \sum_{0 \leq i \leq n} \binom{n}{i} x^i \) (by definition). Therefore, we may write \((1+x)(1+x)^{n-1} = \)

\[
(1+x) \sum_{0 \leq i \leq n} \binom{n-1}{i} x^i = \sum_{0 \leq i \leq n} \binom{n-1}{i} (x^i + x^{i+1}) = \sum_{0 \leq i \leq n} \binom{n-1}{i} x^i + \sum_{0 \leq i \leq n} \binom{n-1}{i} x^{i+1}
\]

The last sum may be written (we take \( j = i + 1 \)):

\[
\sum_{1 \leq j \leq n-1} \binom{n-1}{j-1} x^j = \sum_{0 \leq i \leq n} \binom{n-1}{i-1} x^i,
\]

since the added terms make no contribution. Therefore we have proved that:

\[
\sum_{0 \leq i \leq n} \binom{n}{i} x^i = \sum_{0 \leq i \leq n} \binom{n-1}{i} x^i + \sum_{0 \leq i \leq n} \binom{n-1}{i-1} x^i
\]

which implies the desired result. This is a particular case of the many uses of generating functions we shall see later. The principle is that if two functions are equal – not in a particular point, but across their domain – then their power series developments there coincide.

The result shown above is also true if \( n \) is not a natural number, which we demonstrate via a different proof.

Let us prove that:

\[
\binom{\delta}{i} = \binom{\delta-1}{i} + \binom{\delta-1}{i-1} \quad \delta \text{ is any number}
\]

proof: We have to prove that

\[
\frac{\delta^i}{i!} = \frac{(\delta-1)^i}{i!} + \frac{(\delta-1)^{i-1}}{(i-1)!}
\]
The right-hand side may be written this way:

\[
\frac{1}{(i-1)!} \left( \frac{1}{i} (\delta - 1)(\delta - 2) \ldots (\delta - i) + (\delta - 1)(\delta - 2) \ldots (\delta - i + 1) \right)
\]

\[
= \frac{(\delta - 1)i-1}{(i-1)!} \left( \frac{\delta - i + 1}{i} \right) = \frac{\delta - i + 1}{i} = \binom{\delta}{i}
\]

\[
\binom{r}{m} \binom{m}{k} = \binom{r}{k} \binom{r - k}{m - k} \quad m, k \in \mathbb{N} \quad (*)
\]

Proof: we have to prove that

\[
\frac{r^m m^k}{m! k!} = \frac{r^k (r - k)^{m-k}}{k! (m-k)!}
\]

Notice that \(a^i \cdot a^j = a^{i+j}\) has the "corresponding" rule:

\(r^m = r(r - 1) \ldots (r - k + 1) \cdots (r - m + 1) = r^k (r - k)^{m-k}\). Therefore, we have just shown that \(\binom{r^k (r - k)^{m-k}}{m! k!}
\]

We can now go back to (*) and use it in the right-hand side:

\[
\frac{r^k (r - k)^{m-k}}{k! (m-k)!} = \frac{r^m m^k}{m! k!}
\]

### Summations

- Let \(S_1 \equiv \sum_{i \leq j \leq t} \binom{t}{i} \binom{j}{r}\)

We can drop the range specifier for \(j\) and sum over all of \(\mathbb{N}\) (or even all of \(\mathbb{Z}\), which includes all negative integers as well), since no value outside the specified range makes any contribution!

\[
S_1 = \sum_{i \leq j \leq t} \binom{t}{i} \binom{j - r}{r} = \binom{t}{r} \sum_{k} \binom{t - r}{k} \cdot 1^k = \binom{t}{r} \cdot 2^{-r}
\]

(We have used the binomial theorem in the last step).

- **Vandermonde convolution**

\[
S_2 = \sum_i \binom{p}{i} \binom{d}{n - i} = \binom{p + d}{n}
\]

We provide several proofs of this basic formula:  

**Proof 1:** \((1 + x)^{p+d} = (1 + x)^p (1 + x)^d\), therefore

\[
\sum_j \binom{p+d}{j} x^j = \sum_k \binom{p}{k} x^k \cdot \sum_l \binom{d}{l} x^l
\]
We define a variable $m = k - l$ and we substitute for $l$ in the right member, obtaining the result:

$$\sum_k \binom{p}{k} x^k \cdot \sum_m \binom{d}{m-k} x^{m-k} = \sum_m \left( \sum_k \binom{p}{k} \binom{d}{m-k} \right) x^m$$

Which implies the claim.

**Proof 2:** We will use induction by $p$, arbitrary $d$ (n – arbitrary integer)

basis: $p = 0 \quad \sum_i \binom{0}{i} \binom{d}{n-i} = \binom{d}{n}$

Now we assume this is true up to $p - 1$, in particular, we have the hypothesis: $\sum_i \binom{p-1}{i} \binom{d}{n-i} = \binom{p+d-1}{n-i}$

Now we need to compute

$$\sum_i \binom{p}{i} \binom{d}{n-i} = \sum_i \binom{p-1}{i} \binom{d}{n-i} + \sum_i \binom{p-1}{i-1} \binom{d}{n-i}$$

We have used the recursion transformation for $\binom{n}{i}$.

Now we substitute $k = i - 1$ in the second sum – and it becomes $\sum_k \binom{p-1}{k} \binom{d}{n-k}$. Hence we obtain

$$\binom{p+d-1}{n} + \binom{p+d-1}{n-1} \quad \text{which is} \quad \binom{p+d}{n}$$

**Proof 3:** This proof uses a more “intuitive” method, using the combinatorial interpretation of binomial coefficients:

$$\binom{n}{i} \quad \text{is the number of ways to pick k items out of a set of n distinct items.}$$

Suppose we have $p$ pennies and $d$ dimes (and they are all distinguishable), we may ask the following question: “In how many way can someone pick $n$ out of these coins?”.

The answer is $\binom{p+d}{n}$, but let us obtain it another way:

First count the number of ways using no pennies $\binom{p}{0} \binom{d}{n}$

then we count the number of ways using 1 penny $\binom{p}{1} \binom{d}{n-1}$

... then we count the number of ways using $i$ pennies $\binom{p}{i} \binom{d}{n-i}$

... up to $n$. This way we have proved that $\binom{p+d}{n} = \sum_i \binom{p}{i} \binom{d}{n-i}$.

**Approximation**

Stirling approximation of the factorial ...

$$n! = n^n \sqrt{2\pi n} e^{-n} \left( 1 + \frac{1}{12n} + \frac{1}{288n^2} + \ldots \right)$$
... leads to:

\[
\binom{n}{k} \approx \frac{n!}{k!(n-k)!} = \frac{n^n \sqrt{2\pi n} e^{-n} \cdot n^{k} \cdot n^{n-k}}{k^{k} \sqrt{2\pi k} e^{-k} \sqrt{2\pi (n-k)} e^{-(n-k)}}
\]

\[= \sqrt{\frac{n}{2\pi(n-k)}} \left( \frac{n}{k} \right)^{k} \left( \frac{n}{n-k} \right)^{n-k} \left( 1 + \frac{1}{12} \left( \frac{1}{k} - \frac{1}{n-k} \right) \ldots \right) \]

**Probability**

Usually when we are talking about probability we are using terms like: *sample space* (a common notation for it is \(\Omega\)), *experiments* or *trials*. Let us see what these notions mean, looking at an example.

**Example**

Supposing we have a die. Then \(\Omega = \{1, 2, 3, 4, 5, 6\}\).

Every elementary experiment yields an element of \(\Omega\).

Now we can define \(P\) – a probability over \(\Omega\). We say: every \(e \in \Omega\) has the probability \(P(e)\). For example \(P(3) = \frac{1}{6}\), assuming we have a *perfect* (also called *true*) die. This assumption means that events are *equally likely* (similar terms which are often used: *at random* or *uniform distribution*).

The term *random* should not be used in this sense. Use random as the opposite of deterministic.

An *elementary event* is a possible result of an experiment, while an *event* is a subset of \(\Omega\) – and it may include many elementary events. This means that if \(\omega\) is an event, then its probability

\[
P(\omega) = \sum_{e \in \omega} P(e)
\]

(\(e\) are all the elementary events "contained" in \(\omega\)). In particular:

\[P(\Omega) = 1\]

Now we can introduce an additional concept: *dependence between events*.

Let us go back to the die example. We make some notations:

- \(\omega_e = \text{an even number} \Rightarrow p(\omega_e) = p_2 + p_4 + p_6 = \frac{1}{2}\)
- \(\omega_b = \text{a big number} (\geq 4) \Rightarrow p(\omega_b) = \frac{1}{2}\)

We say that "an event happened" in a trial, if the result of the trial—always an elementary event—is contained in the event. Now we can ask “What is the probability of \(\omega_e\) if we know that \(\omega_b\) happened?”

In the current scenario, where all elementary events are equally likely, we can write:

\[
p(\omega_e | \omega_b) = \frac{\text{number of } \omega_e \text{ elements in } \omega_b}{\text{number of elements in } \omega_b} = \frac{2}{3}
\]

5
In general, if $A$ and $B$ are two events, then

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

$P(A,B) \equiv \text{the probability that both } A \text{ and } B \text{ happen.}$

**The Theorem of Total Probability**

This is one of our few basic computational tools. Assuming we are given $\Omega$ and a partition $\{A_i\}$ of it.

This means that the $A_i$ are disjoint and $\cup A_i = \Omega$. Then

$$P(B) = \sum_{\text{all sets}} P(A_i)P(B|A_i)$$

Here is an example of how this theorem can be used in a calculation:

We have 3 drawers and each drawer contains 2 coins. There are 2 types of coins: golden (G) and silver (S), disposed in the drawers in this way: GG GS SS (but we do not know what is in each drawer). The experiment consists of the following steps:
— choose a drawer,
— take a coin and put it away,
— look at the other coin.

Question: What is the probability that you see G ?

Solution: Using the theorem of total probability, we can write $P(G)$ as a summation:

$$P(G) = P(1)P(G|1) + P(2)P(G|2) + P(3)P(G|3) =$$

$$= \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0 = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

where we assumed that each drawer is equally likely to be opened, and each of the two coins in it equally likely to be taken out first.

Question: What is the probability that you see G if the coin first removed was G ?

$P(G|\text{removed is } G) = P(1)P(G|1, G) + P(2)P(G|2, G) + P(3)P(G|3, G) = \frac{1}{3}$

We have used the notation $P(k, G) \equiv I \text{ chose drawer } k \text{ and found } G.$