Binomial Coefficients

In how many ways can we choose \( k \) elements from an \( n \) element set? There are \( n \) choices for the first element, \( n - 1 \) for the second, ..., down to \( n-k +1 \) for the \( k^{th} \), yielding \( n*(n-1)*...*(n-k+1) \). So there are \( 4*3=12 \) ways to choose 2 elements from 4. Consider \{a,b,c,d\}.\{a,c\} \{a,d\} \{b,c\} \{b,d\} \{c,d\}  

Problem: The above formula counts \((a,b)\) and \((b,a)\). Must divide by \( k! \). In general,  

\[
\binom{r}{k} = \begin{cases} 
\frac{r(r-1)...(r-k+1)}{k(k-1)...1}, & \text{integer } k \geq 0 \\
0, & \text{integer } k < 0
\end{cases}
\]

for arbitrary real (even complex) \( r \).

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \text{integers } n \geq k \geq 0
\]

Note that \( 0! = \prod_{1 \leq k \leq 0} k = 1 \)

Let's look at some values (Pascal's Triangle)

\[
\begin{array}{ccccccc}
0 & 1 & & & & & \\
1 & 1 & 1 & & & & \\
2 & 1 & 2 & 1 & & & \\
3 & 1 & 3 & 3 & 1 & & \\
4 & 1 & 4 & 6 & 4 & 1 & \\
5 & 1 & 5 & 10 & 10 & 5 & 1 \\
6 & 1 & 6 & 15 & 20 & 15 & 6 & \sum=64
\end{array}
\]

Note first 3 columns: \( \binom{n}{0}=1 \) \( \binom{n}{1}=n \) \( \binom{n}{2}=\frac{n(n-1)}{2} \) (triang)

Note symmetry in each row.

\[
\binom{n}{k} = \binom{n}{n-k} \quad \text{integer } n \geq 0, \text{ integer } k \quad \text{Symmetry}
\]

Argument: To choose \( k \) of \( n \) objects, we must choose which \( n-k \) to omit.  

\[
\frac{n!}{k!(n-k)!} \quad \text{same if } k \text{ replaced by } n-k
\]

Sometimes we want to add or get rid of a factor of \( \binom{n}{k} \). Note that  

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n}{k} \left( \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} \right), \quad \text{so}
\]

\[
\binom{r}{k} = \frac{r(r-1)}{k(k-1)} \quad \text{integer } k \neq 0 \quad \text{Absorption/extraction}
\]
From Pascal's triangle, each element = element up + element up-&-left

\[ \binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1} \quad \text{integer } k \]

Combinatorial argument: To choose \( k \) elements from \( n \), mark one of the \( n \)
- it's not included in the \( k \) \( \binom{r-1}{k} \) (choose \( k \) from the rest), or
- it's included in the \( k \) \( \binom{r-1}{k-1} \) (choose \( k-1 \) from the rest)

Applying this several times (keeping constant gap between top & bottom):

always expand smaller lower index

\[
\begin{align*}
\binom{5}{3} &= \binom{4}{3} + \binom{4}{2} = \binom{4}{3} + \binom{3}{2} + \binom{3}{1} = \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{2}{0} \\
&= \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{1}{0} + \binom{1}{1}
\end{align*}
\]

Last term & all subsequent terms are 0. In general, (above \( n=3, r=1 \))

\[
\binom{n+r+1}{n} = \sum_{k \leq n} \binom{r+k}{k} \quad \text{integer } n
\]

Combinatorial interpretation:

Consider all paths from (0 0) -> \( (n \ r+1) \). There are \( n+r+1 \) steps from \{right, up\}, and
\( \binom{n+r+1}{n} \) ways to choose right from them. All paths hit the top line somewhere. Call it \( k, 0 \leq k \leq n \). For each value of \( k \), there are \( \binom{r+k}{k} \) paths first hitting the top line at \( k \).

Repeat the above expansion but, instead of keeping a constant gap between top & bottom,
\[
\binom{5}{3} = \binom{4}{3} + \binom{4}{2} = \binom{3}{3} + \binom{3}{2} + \binom{4}{2} = \binom{2}{3} + \binom{2}{2} + \binom{4}{2} = \binom{1}{3} + \binom{1}{2} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2} = \binom{0}{3} + \binom{1}{2} + \binom{3}{2} + \binom{4}{2} + \binom{4}{2} = \binom{0}{3} + \binom{1}{2} + \binom{3}{2} + \binom{4}{2}
\]

Note that \(\binom{0}{3} = 0\), yielding:

\[
\sum_{0 \leq k \leq n} \binom{k}{m} \quad \text{integers } m, n \geq 0
\]

Combinatorial interpretation: To choose \(m+1\) tickets from a set of \(n+1\) tickets numbered \(0..n\), there are \(\binom{k}{m}\) ways to do this when largest ticket is number \(k\) (choose the remaining \(m\) from \(0..k-1\)).

We often want to manipulate products of binomial coefficients:

\[
\binom{r}{m} \cdot \binom{m}{k} = \binom{r}{k} \cdot \binom{r-k}{m-k} \quad \text{integers } m, k
\]

Combinatorial interpretation:
left side - given \(r\) people, choose \(m\) to play a game, then choose \(k\) of the \(m\) to play offense
right side - given \(r\) people, choose \(k\) to play offense, then from the remaining \(r-k\) choose \(m-k\) to play defense

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

Whence binomial coefficients?

\[
(x + y)^0 = 1 \quad x^0 y^0
\]
\[
(x + y)^1 = 1x + 1y \quad 0 + 1x^1 y^0
\]
\[
(x + y)^2 = 1x^2 + 2x^1 y^1 + 1x^0 y^2
\]

... p What is coefficient of \(x^k y^{n-k}\) in \((x + y)^n\)? In how many ways can we choose \(k\) \(x\)'s & \(n-k\) \(y\)'s (multiplication being commutative, we can gather and accumulate them).

Surprise answer: \(\binom{n}{k}\)
\[
\sum_{k} \binom{r}{k} x^k y^{r-k} = (x + y)^r \quad \text{integers } r \geq 0 \text{ or } |x/y| < 1 \quad \text{Binomial Theorem}
\]

Note: if \(x=y=1\) & \(r=n > 0\) integral, then \(\sum_{k} \binom{n}{k} = 2^n\).

- Also note the sums of the rows of Pascal's triangle.

- Also note there are \(2^n\) \(n\)-bit words, for \(0 \leq k \leq n\), there are \(\binom{n}{k}\) ways to choose the 1-bits in an \(n\)-bit word with \(k\) 1-bits.

- To compute \(\sqrt{1+x}, -1 < x < 1\). Consider \(x=0.04\),

\[
\begin{align*}
\sqrt{1.04} &= (1 + 0.04)^{0.5} = 1 + \frac{1}{2} 0.04 + \frac{1}{2} \frac{(1/2 - 1)}{2} 0.04^2 + \frac{1}{2} \frac{(1/2 - 1)}{2} \frac{(1/2 - 2)}{3} 0.04^3 + \ldots \\
&= 1 + \frac{0.04}{2} - \frac{0.04^2}{8} + \frac{0.04^3}{16} - \ldots = 1 + 0.02 - 0.0002 + 0.000004 - \ldots = 1.019804
\end{align*}
\]

Bounds on binomial coefficients: For \(1 \leq k \leq n\),

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

From Stirling's approximation, \(n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n\), more accurately,

\[
n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \left(1 + \frac{1}{12n} + \frac{1}{22n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \frac{163879}{209018880n^5} + O\left(\frac{1}{n^6}\right)\right)
\]

\[
\begin{align*}
\frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} &= \left(\frac{n}{e}\right) \left(1 + \frac{1}{12n} + \frac{1}{22n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \frac{163879}{209018880n^5} + O\left(\frac{1}{n^6}\right)\right) \\
&\approx \left(\frac{n}{e}\right) \left(1 + \frac{1}{6n} + \frac{1}{8n^2} - \frac{13}{90n^3} - \frac{1}{18n^4} - \frac{1}{10n^5} + O\left(\frac{1}{n^6}\right)\right)
\end{align*}
\]

\[
\begin{align*}
\frac{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k}{k!} &\leq k! \leq \frac{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k}{k!} \left(\frac{1}{12k}\right)^{k+1} &\text{we get that } k! \geq \left(\frac{k}{e}\right)^k, \text{ so} \\
\binom{n}{k} &= \frac{n(n-1)(n-k+1)}{k!} \leq \frac{n^k}{k!} \leq \left(\frac{en}{k}\right)^k &1 \leq k \leq n
\end{align*}
\]

Ex: Consider Direct Chaining Hashing. Assume there are \(m\) buckets (lists), and \(n\) elements distributed among the lists. What is \(E[A']\)?

\[
E[A'] = \sum_{0 \leq k} k \cdot \Pr\{h(key) \text{ has } k \text{ items}\}
\]

\[
\Pr\{1 \text{st } k \text{ items go to } h(key)\} = \left(\frac{1}{m}\right)^k
\]
Pr{last \( n-k \) items somewhere besides \( h(key) \)} = \( \left( \frac{m-1}{m} \right)^{n-k} \)

Pr{\( h(key) \) has exactly \( k \) items} = \( \binom{n}{k} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k} \)

(Note: for \( k=0 \), \( \left( \frac{m-1}{m} \right)^n \); for \( k=n \), \( \left( \frac{1}{m} \right)^n \))

\[ E[A'] = \sum_{0 \leq k} k \binom{n}{k} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k} \]

Remove \( k \) and we can use the binomial theorem, but replacing it with \( m \) or \( n \) means we can move it out of the summation.

Absorption/Extraction:

\[ \binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1} \]

\[ k \binom{n}{k} = n \binom{n-1}{k-1} \]

\[ E[A'] = \sum_{0 \leq k} k \binom{n}{k} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k} = n \sum_{0 \leq k} \binom{n-1}{k-1} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k} \]

2 problems:

- negative term in bottom of \( \binom{n-1}{k-1} \)

- sum of exponents \( \neq n-1 \)

Replace \( k \) by \( k+1 \)

\[ E[A'] = n \sum_{0 \leq k} \binom{n-1}{k} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k} = n \sum_{-1 \leq k} \binom{n-1}{k} \left( \frac{1}{m} \right)^{k+1} \left( \frac{m-1}{m} \right)^{n-k-1} \]

(pull out \( \frac{1}{m} \))

\[ = \frac{n}{m} \sum_{-1 \leq k} \binom{n-1}{k} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k-1} \]

(since \( \binom{n-1}{-1} = 0 \))

\[ = \frac{n}{m} \sum_{0 \leq k} \binom{n-1}{k} \left( \frac{1}{m} \right)^k \left( \frac{m-1}{m} \right)^{n-k-1} \]

(binomial theorem)

\[ = \frac{n}{m} \left( \frac{1}{m} + \frac{m-1}{m} \right)^{n-1} = \frac{n}{m} \]

Variance computed in H.W.#4-91

Ex: Uniform probing hashing, Gonnet 3.3.2, pg. 48

\[ E[A_{n'}] = 1 + E[ | \{ \text{collisions} \} | ] = 1 + \sum_{0 \leq k \leq n} k \cdot \Pr\{ C_n = k \} \]

It's easy to compute \( \Pr\{ C_n > k \} \)

\[ \Pr\{ C_n \geq 1 \} = \frac{n}{m} = \alpha \]

\[ \Pr\{ C_n \geq 2 \} = \frac{n(n-1)}{m(m-1)} \]

\[ \Pr\{ C_n \geq i \} = \prod_{0 \leq k \leq i-1} \frac{n-k}{m-k} \text{ for } i \leq n \quad (0 \text{ for } i > n) \]

Theorem: If r.v. \( X \) takes values from \( \{0,1,2,...\} \), then
Expanding the sum,

\[ \sum_{0 \leq k \leq n} \left( \frac{m-k}{m-n} \right) = \sum_{0 \leq k \leq n} \left( \frac{m-k}{m-n} \right) - \left( \frac{m}{m-n} \right) \]

So \( E[A_{n'}] \) = 1 + \( \frac{1}{m} \) \[ \sum_{0 \leq k \leq n} \left( \frac{m-k}{m-n} \right) - \left( \frac{m}{m-n} \right) \]

Expanding the sum,

\[ \sum_{0 \leq k \leq n} \left( \frac{m-k}{m-n} \right) = \left( \frac{m}{m-n} \right) + \left( \frac{m-1}{m-n} \right) + \ldots + \left( \frac{0}{m-n} \right) = \left( \frac{m+1}{m-n+1} \right) \]

So \( E[A_{n'}] \) = 1 + \( \frac{1}{m} \) \[ \left( \frac{m+1}{m-n+1} \right) - \left( \frac{m}{m-n} \right) \] = 1 + \( \frac{1}{m} \) \[ \left( \frac{(m+1)!}{(m-n+1)!n!} - \frac{m!}{(m-n)!n!} \right) \]

= 1 + \( \frac{1}{m} \) \[ \frac{nm!}{(m-n+1)(m-n)!n!} \] = 1 + \( \frac{1}{m} \) \[ \frac{nm!}{(m-n+1)(m-n)!n!} \]

For large \( m \), \( E[A_{n'}] = \frac{1}{1-\alpha} \) (same as random probing)
**Ex:** Binomial lower bound on worst case number of comparisons for merging. Given two sorted lists of lengths \( n_a \) and \( n_b \), show that under binary decision tree model the number of comparisons necessary to merge them into one sorted list is \( \geq \left\lceil \log \left( \frac{n_a + n_b}{n_a} \right) \right\rceil \).

**Hint:** You may proceed by
- noting that any algorithm whose basic computation is pairwise comparison may be modeled as a binary tree,
- establishing a lower bound on the number of outputs of the algorithm (leaves of the binary decision tree),
- finding the minimum height of a binary tree which has at least a certain number of leaves.

The output of any merge algorithm is a list \( x_1, \ldots, x_m \) where \( m = n_a + n_b \), and the list of \( a \)'s is a sublist of \( x_1, \ldots, x_m \) of length \( n_a \). There are \( \binom{n_a + n_b}{n_a} \) ways to choose this sublist, so the algorithm must be able to distinguish between \( \binom{n_a + n_b}{n_a} \) outputs. The corresponding decision tree must have at least \( \binom{n_a + n_b}{n_a} \) leaves. Any binary tree of \( \geq \binom{n_a + n_b}{n_a} \) leaves must have height \( \geq \left\lceil \log \left( \frac{n_a + n_b}{n_a} \right) \right\rceil \).