

ECS315 2015/1 Part I.2 Dr.Prapun

4 Enumeration / Combinatorics / Counting

There are many probability problems, especially those concerned with gambling, that can ultimately be reduced to questions about cardinalities of various sets. **Combinatorics** is the *study of systematic counting methods*, which we will be using to find the cardinalities of various sets that arise in probability.

4.1 Four Principles

- Addition
- Multiplication
- subtraction }
- Division }

4.1. Addition Principle (Rule of sum):

- When there are m cases such that the i th case has n_i options, for $i = 1, \dots, m$, and no two of the cases have any options in common, the total number of options is $n_1 + n_2 + \dots + n_m$.
- In set-theoretic terms, suppose that a finite set S can be partitioned⁵ into (pairwise disjoint parts) S_1, S_2, \dots, S_m . Then,

$$|S| = |S_1| + |S_2| + \dots + |S_m|.$$



⁵The art of applying the addition principle is to partition the set S to be counted into “manageable parts”; that is, parts which we can readily count. But this statement needs to be qualified. If we partition S into too many parts, then we may have defeated ourselves. For instance, if we partition S into parts each containing only one element, then applying the

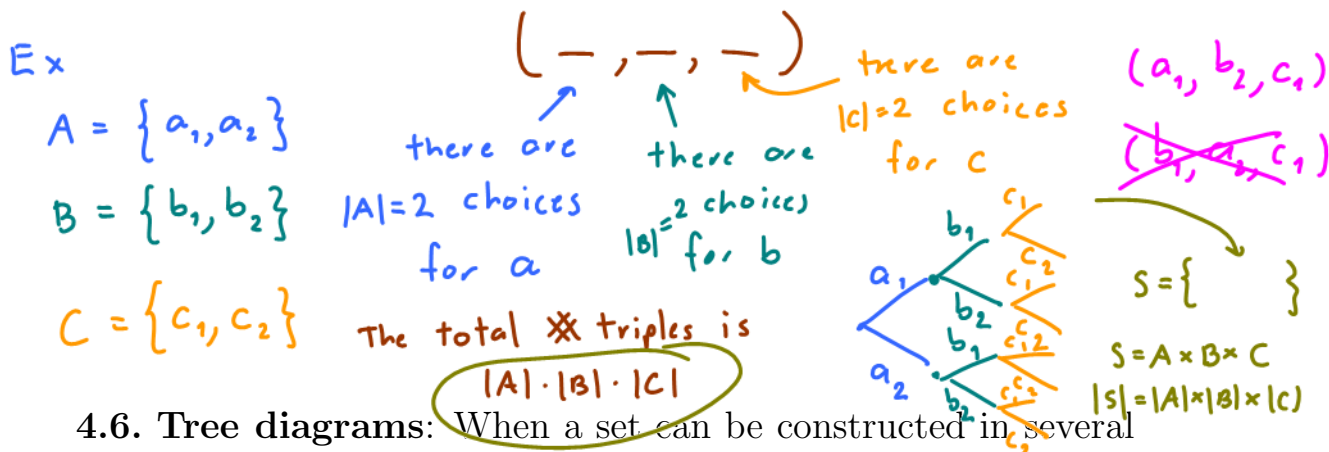
In words, if you can count the number of elements in all of the parts of a partition of S , then $|S|$ is simply the sum of the number of elements in all the parts.

Example 4.2. We may find the number of people living in a country by adding up the number from each province/state.

Example 4.3. [1, p 28] Suppose we wish to find the number of different courses offered by SIIT. We partition the courses according to the department in which they are listed. Provided there is no cross-listing (cross-listing occurs when the same course is listed by more than one department), the number of courses offered by SIIT equals the sum of the number of courses offered by each department.

Example 4.4. [1, p 28] A student wishes to take either a mathematics course or a biology course, but not both. If there are four mathematics courses and three biology courses for which the student has the necessary prerequisites, then the student can choose a course to take in $4 + 3 = 7$ ways.

Example 4.5. Let A , B , and C be finite sets. How many triples are there of the form (a,b,c) , where $a \in A$, $b \in B$, $c \in C$?



4.6. Tree diagrams: When a set can be constructed in several steps or stages, we can represent each of the n_1 ways of completing the first step as a branch of a tree. Each of the ways of completing the second step can be represented as n_2 branches starting from

addition principle is the same as counting the number of parts, and this is basically the same as listing all the objects of S . Thus, a more appropriate description is that the art of applying the addition principle is to partition the set S into not too many manageable parts.[1, p 28]

the ends of the original branches, and so forth. The size of the set then equals the number of branches in the last level of the tree, and this quantity equals

$$n_1 \times n_2 \times \cdots$$

4.7. *Multiplication Principle (Rule of product):*

- When a procedure/operation can be broken down into m **steps**, such that there are n_1 options for step 1, and such that after the completion of step $i - 1$ ($i = 2, \dots, m$) there are n_i options for step i (for each way of completing step $i - 1$), the number of ways of performing the procedure is $n_1 n_2 \cdots n_m$.
- In set-theoretic terms, if sets S_1, S_2, \dots, S_m are finite, then $|S_1 \times S_2 \times \cdots \times S_m| = |S_1| \times |S_2| \times \cdots \times |S_m|$.
- For m finite sets A_1, A_2, \dots, A_m , there are $|A_1| \times |A_2| \times \cdots \times |A_m|$ m -tuples of the form (a_1, a_2, \dots, a_m) where each $a_i \in A_i$.

Example 4.8. Suppose that a deli offers three kinds of bread, three kinds of cheese, four kinds of meat, and two kinds of mustard. How many different meat and cheese sandwiches can you make?

First choose the bread. For each choice of bread, you then have three choices of cheese, which gives a total of $3 \times 3 = 9$ bread/cheese combinations (rye/swiss, rye/provolone, rye/cheddar, wheat/swiss, wheat/provolone ... you get the idea). Then choose among the four kinds of meat, and finally between the two types of mustard or no mustard at all. You get a total of $3 \times 3 \times 4 \times 3 = 108$ different sandwiches.

Suppose that you also have the choice of adding lettuce, tomato, or onion in any combination you want. This choice gives another $2 \times 2 \times 2 = 8$ combinations (you have the choice “yes” or “no” three times) to combine with the previous 108, so the total is now $108 \times 8 = 864$.

That was the multiplication principle. In each step you have several choices, and to get the total number of combinations, multiply. It is fascinating how quickly the number of combinations

grow. Just add one more type of bread, cheese, and meat, respectively, and the number of sandwiches becomes 1,920. It would take years to try them all for lunch. [17, p 33]

Example 4.9 (Slides). In 1961, Raymond Queneau, a French poet and novelist, wrote a book called *One Hundred Thousand Billion Poems*. The book has ten pages, and each page contains a sonnet, which has 14 lines. There are cuts between the lines so that each line can be turned separately, and because all lines have the same rhyme scheme and rhyme sounds, any such combination gives a readable sonnet. The number of sonnets that can be obtained in this way is thus 10^{14} which is indeed a hundred thousand billion. Somebody has calculated that it would take about 200 million years of nonstop reading to get through them all. [17, p 34]

$$\left. \begin{array}{l} 10^x \\ 10^x \\ 10^x \\ \vdots \\ 10^x \\ 10^x \end{array} \right\} 14 \\ \hline = 10^{14}$$

10^2 10^3 10^9

Example 4.10. There are 2^n binary strings/sequences of length n .

Each step has two choices: 0 or 1



2^A 2 possibilities of A

Example 4.11. For a finite set A , the cardinality of its power set 2^A is

$$|2^A| = 2^{|A|}$$

The set of all subsets of A including the empty set and A itself.

$$A = \{a_1, a_2, a_3, \dots, a_{|A|}\}$$

$x \checkmark x$
 $\checkmark x \checkmark$

$$\{a, b, c\} = A$$

$$2^A = \{ \emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\} \}$$

Example 4.12. (Slides) Jack is so busy that he's always throwing his socks into his top drawer without pairing them. One morning Jack oversleeps. In his haste to get ready for school, (and still a bit sleepy), he reaches into his drawer and pulls out 2 socks. Jack knows that 4 blue socks, 3 green socks, and 2 tan socks are in his drawer.

A

(a) What are Jack's chances that he "pulls out 2 blue socks" to match his blue slacks?

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{4}{2}}{\binom{9}{2}} = \frac{1}{6}$$

$$|\Omega| = \frac{9 \times 8}{2}$$

$$|A| = \frac{4 \times 3}{2}$$

(b) What are the chances that he "pulls out a pair of matching socks?"



$$P(S) = \frac{|S|}{|\Omega|} = \frac{|A_b| + |A_g| + |A_t|}{(4 \times 3) + (3 \times 2) + (2 \times 1)} = \frac{5}{18}$$

Example 4.13. [1, p 29–30] Determine the number of positive integers that are factors of the number

$$3^4 \times 5^2 \times 11^7 \times 13^8.$$

The numbers 3, 5, 11, and 13 are prime numbers. By the fundamental theorem of arithmetic, each factor is of the form

$$3^i \times 5^j \times 11^k \times 13^\ell,$$

where $0 \leq i \leq 4$, $0 \leq j \leq 2$, $0 \leq k \leq 7$, and $0 \leq \ell \leq 8$. There are five choices for i , three for j , eight for k , and nine for ℓ . By the multiplication principle, the number of factors is

$$5 \times 3 \times 8 \times 9 = 1080.$$

4.14. Subtraction Principle: Let A be a set and let S be a larger set containing A . Then



$$|A| = |S| - |S \setminus A|$$

- When S is the same as Ω , we have $|A| = |\Omega| - |A^c|$
- Using the subtraction principle makes sense only if it is easier to count the number of objects in S and in $S \setminus A$ than to count the number of objects in A .

Example 4.15. *Chevalier de Mere's Scandal of Arithmetic:*

Which is more likely, obtaining at least one six in 4 tosses of a fair dice (event A), or obtaining at least one double six in 24 tosses of a pair of dice (event B)?

$$P(A) = \frac{|A|}{|\Omega|} = \frac{|\Omega| - |A^c|}{|\Omega|} \quad P(B) = \frac{|B|}{|\Omega|}$$

Handwritten notes: 6^4 under the first fraction, 24 under the second, and 6^c never shows up with an arrow pointing to $|A^c|$. For $P(B)$, 24 is written below the denominator and $(36)^{24}$ is written below the numerator.

We have

$$P(A) = \frac{6^4 - 5^4}{6^4} = 1 - \left(\frac{5}{6}\right)^4 \approx .518$$

and

$$P(B) = \frac{36^{24} - 35^{24}}{36^{24}} = 1 - \left(\frac{35}{36}\right)^{24} \approx .491.$$

Therefore, the first case is more probable.

Remark 1: Probability theory was originally inspired by gambling problems. In 1654, Chevalier de Mere invented a gambling system which bet even money⁶ on event B above. However, when he began losing money, he asked his mathematician friend Pascal to analyze his gambling system. Pascal discovered that the Chevalier's system would lose about 51 percent of the time. Pascal became so interested in probability and together with another famous mathematician, Pierre de Fermat, they laid the foundation of probability theory. [U-X-L Encyclopedia of Science]

Remark 2: de Mere originally claimed to have discovered a *contradiction in arithmetic*. De Mere correctly knew that it was advantageous to wager on occurrence of event A, but his experience as gambler taught him that it was not advantageous to wager on occurrence of event B. He calculated $P(A) = 1/6 + 1/6 + 1/6 + 1/6 = 4/6$ and similarly $P(B) = 24 \times 1/36 = 24/36$ which is the same as $P(A)$. He mistakenly claimed that this evidenced a contradiction to the arithmetic law of proportions, which says that $\frac{4}{6}$ should be the same as $\frac{24}{36}$. Of course we know that he could not simply add up the probabilities from each tosses. (By De Meres logic, the probability of at least one head in two tosses of a fair coin would be $2 \times 0.5 = 1$, which we know cannot be true). [21, p 3]

4.16. Division Principle (Rule of quotient): When a finite set S is partitioned into equal-sized parts of m elements each, there are $\frac{|S|}{m}$ parts. *Ex. Suppose we have a room of 20 students. Quiz is conducted in a group of four students.*

⁶Even money describes a wagering proposition in which if the bettor loses a bet, he or she stands to lose the same amount of money that the winner of the bet would win.

How many groups are there? $\frac{20}{4} = 5$.

4.2 Four Kinds of Counting Problems



4.17. Choosing objects from a collection is called **sampling**, and the chosen objects are known as a **sample**. The four kinds of counting problems are [9, p 34]:

- (a) Ordered sampling of r out of n items with replacement: n^r ;
- (b) Ordered sampling of $r \leq n$ out of n items without replacement: $(n)_r$;
- (c) Unordered sampling of $r \leq n$ out of n items without replacement: $\binom{n}{r}$;
- (d) Unordered sampling of r out of n items with replacement: $\binom{n+r-1}{r}$.

- See 4.33 for “bars and stars” argument.

Many counting problems can be simplified/solved by realizing that they are equivalent to one of these counting problems.

4.18. Ordered Sampling: Given a set of n distinct items/objects, select a distinct **ordered**⁷ sequence (word) of length r drawn from this set.

- (a) **Ordered sampling with replacement:** $\mu_{n,r} = n^r$
 - Ordered sampling of r out of n items with replacement.
 - The “with replacement” part means “an object can be chosen repeatedly.”
 - Example: From a deck of n cards, we draw r cards with replacement; i.e., we draw a card, make a note of it, **put the card back in** the deck and re-shuffle the deck before choosing the next card. How many different sequences of r cards can be drawn in this way? [9, Ex. 1.30]

$$\underbrace{\underbrace{n \times n \times n \times \dots \times n}_{r \text{ times}}}_{r \text{ times}} = n^r$$

⁷Different sequences are distinguished by the order in which we choose objects.

(b) *Ordered sampling without replacement:*

$$\begin{aligned} (n)_r &= \prod_{i=0}^{r-1} (n-i) = \frac{n!}{(n-r)!} \\ &= \underbrace{n \times (n-1) \times \dots \times (n-(r-1))}_{r \text{ terms}}; \quad r \leq n \end{aligned}$$

- Ordered sampling of $r \leq n$ out of n items without replacement.
- The “without replacement” means “once we choose an object, we remove that object from the collection and we cannot choose it again.”
- In Excel, use PERMUT(n, r).
- Sometimes referred to as “the number of possible r -permutations of n distinguishable objects”
- Example: The number of sequences⁸ of size r drawn from an alphabet of size n without replacement.

$(3)_2 = 3 \times 2 = 6$ is the number of sequences of size 2 drawn from an alphabet of size 3 without replacement.

Suppose the alphabet set is $\{A, B, C\}$. We can list all sequences of size 2 drawn from $\{A, B, C\}$ without replacement:

A B
A C
B A
B C
C A
C B

- Example: From a deck of 52 cards, we draw a hand of 5 cards without replacement (drawn cards are not placed back in the deck). How many hands can be drawn in this way?

⁸Elements in a sequence are ordered.

$$(52)_5 = \underline{52} \times \underline{51} \times \underline{50} \times \underline{49} \times \underline{48}$$

- For integers r, n such that $r > n$, we have $(n)_r = 0$.
- We define $(n)_0 = 1$. (This makes sense because we usually take the empty product to be 1.)
- $(n)_1 = n$
- $(n)_r = (n - (r - 1))(n)_{r-1}$. For example, $(7)_5 = (7 - 4)(7)_4$.
- $(1)_r = \begin{cases} 1, & \text{if } r = 1 \\ 0, & \text{if } r > 1 \end{cases}$
- Extended definition: The definition in product form

$$(n)_r = \prod_{i=0}^{r-1} (n - i) = \underbrace{n \cdot (n - 1) \cdots (n - (r - 1))}_{r \text{ terms}}$$

can be extended to *any real number* n and a non-negative integer r .

Example 4.19. (Slides) The Seven Card Hustle: Take five red cards and two black cards from a pack. Ask your friend to shuffle them and then, without looking at the faces, lay them out in a row. Bet that they can't turn over three red cards. The probability that they CAN do it is

Definition 4.20. For any integer n greater than 1, the symbol $n!$, pronounced "***n factorial***," is defined as the **product of all positive integers less than or equal to n .**

$$n! = n \times (n-1) \times (n-2) \times \dots \times 2 \times 1$$

$$4! = \overbrace{4 \times 3}^{12} \times 2 \times 1 = 24$$

(a) $0! = 1! = 1$

(b) $n! = n(n - 1)!$

(c) $n! = \int_0^{\infty} e^{-t} t^n dt$

(d) Computation:

$$52! \approx 8 \times 10^{67}$$

- (i) MATLAB: Use `factorial(n)`. Since double precision numbers only have about 15 digits, the answer is only accurate for $n \leq 21$. For larger n , the answer will have the right magnitude, and is accurate for the first 15 digits.
 - (ii) Google's web search box built-in calculator: Use `n!`
- (e) Approximation: Stirling's Formula [5, p. 52]:

$$n! \approx \sqrt{2\pi n} n^n e^{-n} = \left(\sqrt{2\pi e}\right) e^{(n+\frac{1}{2})\ln(\frac{n}{e})}. \quad (2)$$

In some references, the sign \approx is replaced by \sim to emphasize that the ratio of the two sides converges to unity as $n \rightarrow \infty$.

4.21. Factorial and Permutation: The number of arrangements (permutations) of $n \geq 0$ distinct items is $(n)_n = n!$.

- Meaning: The number of ways that n distinct objects can be ordered.
 - A special case of ordered sampling without replacement where $r = n$.
- In MATLAB, use `perms(v)`, where v is a row vector of length n , to create a matrix whose rows consist of all possible permutations of the n elements of v . (So the matrix will contain $n!$ rows and n columns.)

Example 4.22. In MATLAB, `perms([3 4 7])` gives

7 4 3
7 3 4
4 7 3
4 3 7
3 4 7
3 7 4

$n=3$
 $n! = 3! = 3 \times 2 \times 1 = 6$

`perms([3 3 7])`

still list 6 results:

7 3 3
7 3 3
3 7 3
3 3 7
3 3 7
3 7 3

* elements
* groups = $3!$
* elements / group = $2!$

Similarly, perms($\overbrace{\text{'abcd'}}^{n=4}$) gives

dcba dcab dbca dbac dabc dacb
 cdba cdab cbda cbad cabd cadb
 bcda bead bdca bdac badc bacd
 acbd acdb abcd abdc adbc adcb

$$n! = 4! = 24$$

Example 4.23. (Slides) Finger-Smudge on Touch-Screen Devices

Example 4.24. (Slides) *Probability of coincidence birthday:* Probability that there is at least two people who have the same birthday⁹ in a group of r persons:

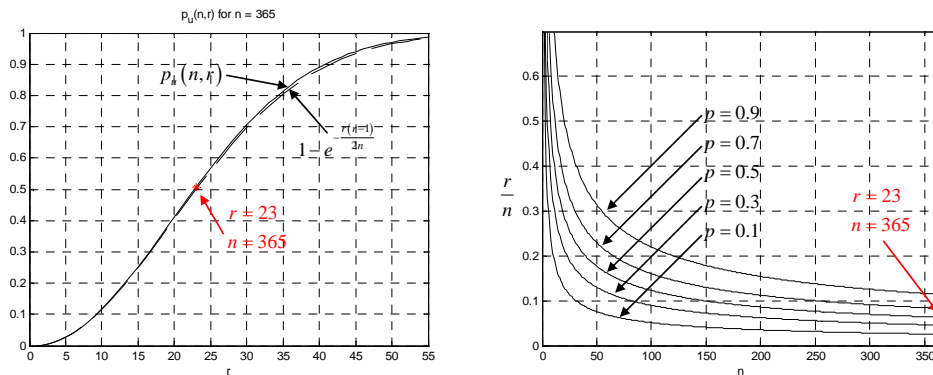


Figure 1: $p_u(n, r)$: The probability of the event that at least one element appears twice in random sample of size r with replacement is taken from a population of n elements.

Example 4.25. It is surprising to see how quickly the probability in Example 4.24 approaches 1 as r grows larger.

⁹We ignore February 29 which only comes in leap years.

Birthday Paradox: In a group of 23 randomly selected people, the probability that at least two will share a birthday (assuming birthdays are equally likely to occur on any given day of the year¹⁰) is about 0.5.

- At first glance it is surprising that the probability of 2 people having the same birthday is so large¹¹, since there are only 23 people compared with 365 days on the calendar. Some of the surprise disappears if you realize that there are $\binom{23}{2} = 253$ pairs of people who are going to compare their birthdays. [3, p. 9]

Example 4.26. Another variant of the birthday coincidence paradox: The group size must be at least 253 people if you want a probability > 0.5 that someone will have the same birthday as you. [3, Ex. 1.13] (The probability is given by $1 - \left(\frac{364}{365}\right)^r$.)

- A naive (but incorrect) guess is that $\lceil 365/2 \rceil = 183$ people will be enough. The “problem” is that many people in the group will have the same birthday, so the number of different birthdays is smaller than the size of the group.
- On late-night television’s *The Tonight Show* with Johnny Carson, Carson was discussing the birthday problem in one of his famous monologues. At a certain point, he remarked to his audience of approximately 100 people: “Great! There must

¹⁰In reality, birthdays are not uniformly distributed. In which case, the probability of a match only becomes larger for any deviation from the uniform distribution. This result can be mathematically proved. Intuitively, you might better understand the result by thinking of a group of people coming from a planet on which people are always born on the same day.

¹¹In other words, it was surprising that the size needed to have 2 people with the same birthday was so small.

be someone here who was born on my birthday!” He was off by a long shot. Carson had confused two distinctly different probability problems: (1) the probability of one person out of a group of 100 people having the same birth date as Carson himself, and (2) the probability of any two or more people out of a group of 101 people having birthdays on the same day. [21, p 76]

4.27. Now, let’s revisit ordered sampling of r out of n different items without replacement. This is also called the number of possible **r -permutations** of n different items. One way to look at the sampling is to first consider the $n!$ permutations of the n items. Now, use only the first r positions. Because we do not care about the last $n - r$ positions, we will group the permutations by the first r positions. The size of each group will be the number of possible permutations of the $n - r$ items that has not already been used in the first r positions. So, each group will contain $(n - r)!$ members. By the division principle, the number of groups is $n!/(n - r)!$.

4.28. The number of permutations of $n = n_1 + n_2 + \dots + n_r$ objects of which n_1 are of one type, n_2 are of the second type, n_3 are of the third type, ..., and n_r are of the r th type is

*n objects
r types*

$$\frac{n!}{n_1!n_2!\dots n_r!} = \binom{n}{n_1, n_2, \dots, n_r}$$

Example 4.29. The number of permutations of A₁A₂BC

AAABC

*A₁BA₂C
A₂BA₁C*

$$\frac{4!}{2!1!1!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 1 \times 1} = 12 \quad \binom{5}{3,1,1} = \frac{5!}{3!1!1!} = 20$$

Example 4.30. Bar Codes: A part is labeled by printing with four thick lines, three medium lines, and two thin lines. If each ordering of the nine lines represents a different label, how many different labels can be generated by using this scheme?

AA BB CCC



$$\frac{9!}{4!3!2!} = 1260$$

4.31. Binomial coefficient:

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

$$\frac{7!}{2!2!3!}$$



$$\binom{5}{2} = \frac{5!}{3!2!} = \frac{5 \times 4 \times \cancel{3!}}{\cancel{3!} \times 2} = 10$$

*A₁BCA₂A₃
A₂BCA₁A₃
A₁BCA₃A₂
A₂BCA₃A₁
A₃BCA₁A₂
A₃BCA₂A₁*

- (a) Read “ n choose r ”.
- (b) Meaning:
- (i) **Unordered sampling** of $r \leq n$ out of n distinct items **without replacement**

$n = 5$
choose

- (ii) The number of subsets of size r that can be formed from a set of n elements (without regard to the order of selection).
- (iii) The number of combinations of n objects selected r at a time.
- (iv) the number of **r -combinations** of n objects.
- (v) The number of (unordered) sets of size r drawn from an alphabet of size n without replacement.
- (c) Computation:
- (i) MATLAB:
- `nchoosek(n,r)`, where n and r are nonnegative integers, returns $\binom{n}{r}$.
 - `nchoosek(v,r)`, where v is a row vector of length n , creates a matrix whose rows consist of all possible combinations of the n elements of v taken r at a time. The matrix will contains $\binom{n}{r}$ rows and r columns.
 - Example: `nchoosek('abcd',2)` gives


```
ab
ac
ad
```

bc
bd
cd

(ii) Excel: `combin(n,r)`

(iii) Mathcad: `combin(n,r)`

(iv) Maple: $\binom{n}{r}$

(v) Google's web search box built-in calculator: `n choose r`

(d) Reflection property: $\binom{n}{r} = \binom{n}{n-r}$.

(e) $\binom{n}{n} = \binom{n}{0} = 1$.

(f) $\binom{n}{1} = \binom{n}{n-1} = n$.

(g) $\binom{n}{r} = 0$ if $n < r$ or r is a negative integer.

(h) $\max_r \binom{n}{r} = \binom{n}{\lfloor \frac{n+1}{2} \rfloor}$.

Example 4.32. In bridge, 52 cards are dealt to four players; hence, each player has 13 cards. The order in which the cards are dealt is not important, just the final 13 cards each player ends up with. How many different bridge games can be dealt? (Answer: 53,644,737,765,488,792,839,237,440,000)

4.33. The bars and stars argument:

- Example: Find all nonnegative integers x_1, x_2, x_3 such that

$$x_1 + x_2 + x_3 = 3.$$

$0 + 0 + 3$	$1\ 1\ 1$
$0 + 1 + 2$	$1\ 1\ 1$
$0 + 2 + 1$	$1\ 1\ 1$
$0 + 3 + 0$	$1\ 1\ 1$
$1 + 0 + 2$	$1\ 1\ 1$
$1 + 1 + 1$	$1\ 1\ 1$
$1 + 2 + 0$	$1\ 1\ 1$
$2 + 0 + 1$	$1\ 1\ 1$
$2 + 1 + 0$	$1\ 1\ 1$
$3 + 0 + 0$	$1\ 1\ 1$

- There are $\binom{n+r-1}{r} = \binom{n+r-1}{n-1}$ distinct n -tuples (x_1, x_2, \dots, x_n) of nonnegative integers such that $x_1 + x_2 + \dots + x_n = r$. We use $n - 1$ bars to separate r 1's.
 - (a) Suppose we further require that the x_i are strictly positive ($x_i \geq 1$), then there are $\binom{r-1}{n-1}$ solutions.
 - (b) **Extra Lower-bound Requirement:** Suppose we further require that $x_i \geq a_i$ where the a_i are some given nonnegative integers, then the number of solution is

$$\binom{r - (a_1 + a_2 + \dots + a_n) + n - 1}{n - 1}.$$

Note that here we work with equivalent problem: $y_1 + y_2 + \dots + y_n = r - \sum_{i=1}^n a_i$ where $y_i \geq 0$.

- Consider the distribution of $r = 10$ indistinguishable balls into $n = 5$ distinguishable cells. Then, we only concern with the number of balls in each cell. Using $n - 1 = 4$ bars, we can divide $r = 10$ stars into $n = 5$ groups. For example, $****|***||**|*$ would mean $(4,3,0,2,1)$. In general, there are $\binom{n+r-1}{r}$ ways of arranging the bars and stars.

4.34. Unordered sampling with replacement: There are n items. We sample r out of these n items with replacement. Because the order in the sequences is not important in this kind of sampling, two samples are distinguished by the number of each item in the sequence. In particular, Suppose r letters are drawn

with replacement from a set $\{a_1, a_2, \dots, a_n\}$. Let x_i be the number of a_i in the drawn sequence. Because we sample r times, we know that, for every sample, $x_1 + x_2 + \dots + x_n = r$ where the x_i are non-negative integers. Hence, there are $\binom{n+r-1}{r}$ possible unordered samples with replacement.

4.3 Binomial Theorem and Multinomial Theorem

4.35. Binomial theorem: Sometimes, the number $\binom{n}{r}$ is called a *binomial coefficient* because it appears as the coefficient of $x^r y^{n-r}$ in the expansion of the binomial $(x+y)^n$. More specifically, for any positive integer n , we have,

$$(x+y)^n = \sum_{r=0}^n \binom{n}{r} x^r y^{n-r} \quad (3)$$

(Slide) To see how we get (3), let's consider a smaller case of $n = 3$. The expansion of $(x+y)^3$ can be found using combinatorial reasoning instead of multiplying the three terms out. When $(x+y)^3 = (x+y)(x+y)(x+y)$ is expanded, all products of a term in the first sum, a term in the second sum, and a term in the third sum are added. Terms of the form x^3 , x^2y , xy^2 , and y^3 arise. To obtain a term of the form x^3 , an x must be chosen in each of the sums, and this can be done in only one way. Thus, the x^3 term in the product has a coefficient of 1. To obtain a term of the form x^2y , an x must be chosen in two of the three sums (and consequently a y in the other sum). Hence, the number of such terms is the number of 2-combinations of three objects, namely, $\binom{3}{2}$. Similarly, the number of terms of the form xy^2 is the number of ways to pick one of the three sums to obtain an x (and consequently take a y from each of the other two terms). This can be done in $\binom{3}{1}$ ways. Finally, the only way to obtain a y^3 term is to choose the y for each of the three sums in the product, and this can be done in exactly one way. Consequently, it follows that

$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3.$$

Now, let's state a combinatorial proof of the binomial theorem (3). The terms in the product when it is expanded are of the form

$x^r y^{n-r}$ for $r = 0, 1, 2, \dots, n$. To count the number of terms of the form $x^r y^{n-r}$, note that to obtain such a term it is necessary to choose r x s from the n sums (so that the other $n - r$ terms in the product are y s). Therefore, the coefficient of $x^r y^{n-r}$ is $\binom{n}{r}$.

From (3), if we let $x = y = 1$, then we get another important identity:

$$\sum_{r=0}^n \binom{n}{r} = 2^n. \quad (4)$$

4.36. Multinomial Counting: The multinomial coefficient

$$\binom{n}{n_1 \ n_2 \ \cdots \ n_r}$$

is defined as

$$\begin{aligned} \prod_{i=1}^r \binom{n - \sum_{k=0}^{i-1} n_k}{n_i} &= \binom{n}{n_1} \cdot \binom{n - n_1}{n_2} \cdot \binom{n - n_1 - n_2}{n_3} \cdots \binom{n_r}{n_r} \\ &= \frac{n!}{\prod_{i=1}^r n_i!}. \end{aligned}$$

We have seen this before in (4.28). It is the number of ways that we can arrange $n = \sum_{i=1}^r n_i$ tokens when having r types of symbols and n_i indistinguishable copies/tokens of a type i symbol.

4.37. Multinomial Theorem:

$$(x_1 + \dots + x_r)^n = \sum \frac{n!}{i_1! i_2! \cdots i_r!} x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r},$$

where the sum ranges over all ordered r -tuples of integers i_1, \dots, i_r satisfying the following conditions:

$$i_1 \geq 0, \dots, i_r \geq 0, \quad i_1 + i_2 + \cdots + i_r = n.$$

When $r = 2$ this reduces to the binomial theorem.