CMSC 207- Lecture 25
CHAPTER 9: Counting and Probability (9.3 & 9.4)

Dr. Ahmed Tarek

Counting Elements of Disjoint Sets: The Addition Rule

This rule states that the number of elements in a union of mutually disjoint finite sets equals the sum of the number of elements in each of the component sets.

Theorem 9.3.1 The Addition Rule

Suppose a finite set A equals the union of k distinct mutually disjoint subsets A_1, $A_2, \ldots, A_k$. Then

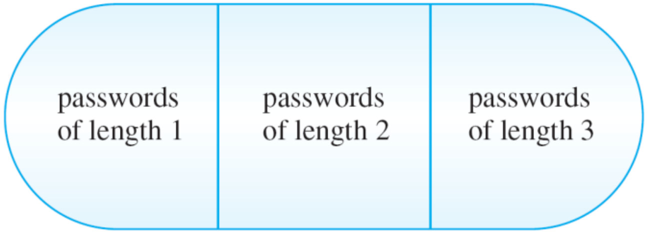
$$N(A) = N(A_1) + N(A_2) + \cdots + N(A_k).$$

Example 1 – *Counting Passwords with Three or Fewer Letters*

A computer access password consists of from one to three letters chosen from the 26 in the alphabet with repetitions allowed. How many different passwords are possible?

Solution:

The set of all passwords can be partitioned into subsets consisting of those of length 1, those of length 2, and those of length 3 as shown in Figure.



Set of All Passwords of Length ≤ 3

Example 1 – *Solution*

By the addition rule, the total number of passwords equals the number of passwords of length 1, plus the number of passwords of length 2, plus the number of passwords of length 3. Now the

number of passwords of length $1 = 26$ because there are 26 letters in the alphabet

number of passwords of length $2 = 26^2$ because forming such a word can be thought of as a two-step process in which there are 26 ways to perform each step

Example 1 – *Solution*

number of passwords of length 3 $= 26^3$

because forming such a word can be thought of as a three-step process in which there are 26 ways to perform each step.

Hence the total number of passwords = $26 + 26^2 + 26^3$

= 18,278.

The Difference Rule

An important consequence of the addition rule is the fact that if the number of elements in a set *A* and the number in a subset *B* of *A* are both known, then the number of elements that are in *A* and not in *B* can be computed.

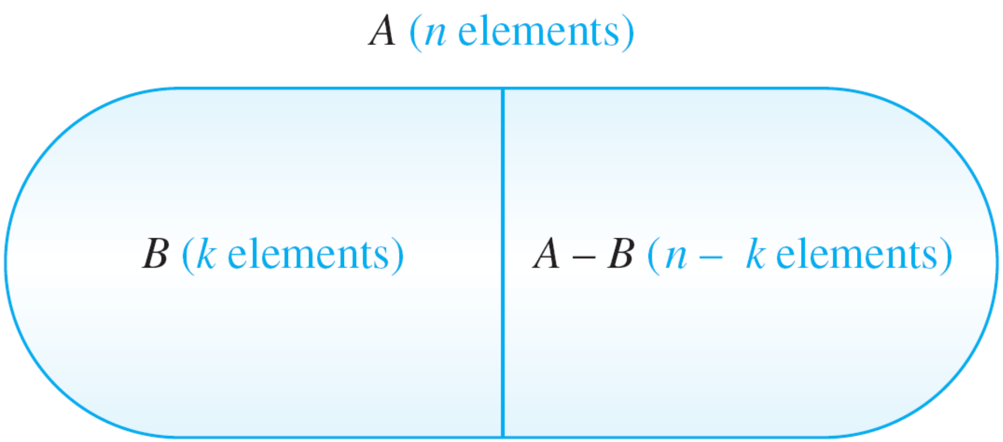
Theorem 9.3.2 The Difference Rule

If A is a finite set and B is a subset of A, then

$$N(A - B) = N(A) - N(B).$$

The Difference Rule

The difference rule is illustrated in Figure.



The Difference Rule

The Difference Rule

The difference rule holds for the following reason: If B is a subset of A, then the two sets B and $A - B$ have no elements in common and $B \cup (A - B) = A$. Hence, by the addition rule,

$$N(B) + N(A - B) = N(A).$$

Subtracting $N(B)$ from both sides gives the equation: $N(A - B) = N(A) - N(B)$.

Example 3 – *Counting PINs with Repeated Symbols*

A typical PIN (personal identification number) is a sequence of any four symbols chosen from the 26 letters in the alphabet and the ten digits, with repetition allowed.

a. How many PINs contain repeated symbols?
b. If all PINs are equally likely, what is the probability that a randomly chosen PIN contains a repeated symbol?

Example 3(a) – *Solution*

There are $36^4 = 1{,}679{,}616$ PINs when repetition is allowed, and there are $36 \cdot 35 \cdot 34 \cdot 33 = 1{,}413{,}720$ PINs when repetition is not allowed.

Thus, by the difference rule, there are

$$1{,}679{,}616 - 1{,}413{,}720 = 265{,}896$$

PINs that contain at least one repeated symbol.

Example 3(b) – *Solution*

There are 1,679,616 PINs in all, and by part (**a**) 265,896 of these contain at least one repeated symbol.

Thus, by the equally likely probability formula, the probability that a randomly chosen PIN contains a repeated symbol is

$$\frac{265{,}896}{1{,}679{,}616} \cong 0.158 = 15.8\%.$$

The Difference Rule

An alternative solution to Example 3(**b**) is based on the observation that if *S* is the set of all PINs and *A* is the set of all PINs with no repeated symbol, then $S - A$ is the set of all PINs with at least one repeated symbol. It follows that

$$
\begin{aligned}
P(S-A) &= \frac{N(S-A)}{N(S)} && \text{by definition of probability in the equally likely case} \\
&= \frac{N(S)-N(A)}{N(S)} && \text{by the difference rule} \\
&= \frac{N(S)}{N(S)} - \frac{N(A)}{N(S)} && \text{by the laws of fractions} \\
&= 1 - P(A) && \text{by definition of probability in the equally likely case}
\end{aligned}
$$

The Difference Rule

We know that the probability that a PIN chosen at random contains no repeated symbol is $P(A)$

$= \dfrac{1{,}413{,}720}{1{,}679{,}616} \cong .8417$

And hence

$$
\begin{aligned}
P(S - A) &\cong 1 - 0.842 \\
&\cong 0.158 \\
&= 15.8\%
\end{aligned}
$$

The Difference Rule

This solution illustrates a more general property of probabilities: that the probability of the complement of an event is obtained by subtracting the probability of the event from the number 1.

Formula for the Probability of the Complement of an Event

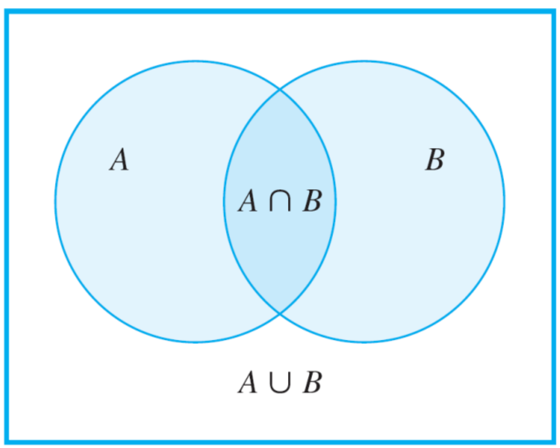
If S is a finite sample space and A is an event in S, then

$$P(A^c) = 1 - P(A).$$

The Inclusion/Exclusion Rule

The addition rule says how many elements are in a union of sets if the sets are mutually disjoint. Now consider the question of how to determine the number of elements in a union of sets when some of the sets overlap.

or simplicity, begin by looking at a union of two sets *A* and *B*, as shown in Figure.



The Inclusion/Exclusion Rule

To get an accurate count of the elements in $A \cup B$, it is necessary to subtract the number of elements that are in both A and B. Because these are the elements in $A \cap B$, $N(A \cup B) = N(A) + N(B) - N(A \cap B)$.

A similar analysis gives a formula for the number of elements in a union of three sets, as shown in Theorem 9.3.3.

Theorem 9.3.3 The Inclusion/Exclusion Rule for Two or Three Sets

If A, B, and C are any finite sets, then

$$N(A \cup B) = N(A) + N(B) - N(A \cap B)$$

and

$$N(A \cup B \cup C) = N(A) + N(B) + N(C) - N(A \cap B) - N(A \cap C) - N(B \cap C) + N(A \cap B \cap C).$$

Example 6 – *Counting Elements of a General Union*

a. How many integers from 1 through 1,000 are multiples of 3 or multiples of 5?
b. How many integers from 1 through 1,000 are neither multiples of 3 nor multiples of 5?

Solution:

a. Let A = the set of all integers from 1 through 1,000 that are multiples of 3.

Let B = the set of all integers from 1 through 1,000 that are multiples of 5.

Example 6 – *Solution*

Then $A \cup B$ = the set of all integers from 1 through 1,000 that are multiples of 3 or multiples of 5
and

$A \cap B$ = the set of all integers from 1 through 1,000 that are multiples of both 3 and 5
= the set of all integers from 1 through 1,000 that are multiples of 15.

Example 6 – *Solution*

Because every third integer from 3 through 999 is a multiple of 3, each can be represented in the form $3k$, for some integer k from 1 through 333.

Hence there are 333 multiples of 3 from 1 through 1,000, and so $N(A) = 333$.

$$\boxed{\begin{array}{ccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & \ldots & 996 & 997 & 998 & 999 \\ & & \updownarrow & & & \updownarrow & & \updownarrow & & & \updownarrow \\ & & 3\cdot 1 & & & 3\cdot 2 & & 3\cdot 332 & & & 3\cdot 333 \end{array}}$$

Example 6 – *Solution*

Similarly, each multiple of 5 from 1 through 1,000 has the form 5*k*, for some integer *k* from 1 through 200.

$$
\begin{array}{cccccccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \ldots & 995 & 996 & 997 & 998 & 999 & 1{,}000 \\
 & & & & \updownarrow & & & & & \updownarrow & & \updownarrow & & & & & \updownarrow \\
 & & & & 5 \cdot 1 & & & & & 5 \cdot 2 & & 5 \cdot 199 & & & & & 5 \cdot 200
\end{array}
$$

Thus there are 200 multiples of 5 from 1 through 1,000 and $N(B) = 200$.

Example 6 – *Solution*

Finally, each multiple of 15 from 1 through 1,000 has the form 15*k*, for some integer *k* from 1 through 66 (since $990 = 66 \cdot 15$).

1	2	...	15	...	30	...	975	...	990	...	999	1,000
			$\updownarrow$		$\updownarrow$		$\updownarrow$		$\updownarrow$			
			$15 \cdot 1$		$15 \cdot 2$		$15 \cdot 65$		$15 \cdot 66$			

Hence there are 66 multiples of 15 from 1 through 1,000, and $N(A \cap B) = 66$.

Example 6 – *Solution*

It follows by the inclusion/exclusion rule that

$$\begin{aligned} N(A \cup B) &= N(A) + N(B) - N(A \cap B) \\ &= 333 + 200 - 66 \\ &= 467. \end{aligned}$$

Thus, 467 integers from 1 through 1,000 are multiples of 3 or multiples of 5.

Example 6 – *Solution*

b. There are 1,000 integers from 1 through 1,000, and by part (**a**), 467 of these are multiples of 3 or multiples of 5.

Thus, by the set difference rule, there are 1,000 – 467 = 533 that are neither multiples of 3 nor multiples of 5.

In-class Assignment #1

How many strings of hexadecimal digits consist of from one through three digits? (Hexadecimal numbers are constructed using the 16 digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F)?

How many strings of hexadecimal digits consist of from two through four digits?

The Inclusion/Exclusion Rule

Note that the solution to part (**b**) of Example 6 hid a use of De Morgan's law.
The number of elements that are neither in A nor in B is $N(A^c \cap B^c)$, and by De Morgan's law, $A^c \cap B^c = (A \cup B)^c$.

So $N((A \cup B)^c)$ was then calculated using the set difference rule: $N((A \cup B)^c) = N(U) - N(A \cup B)$, where the universe U was the set of all integers from 1 through 1,000.

The Pigeonhole Principle

The pigeonhole principle states that if n pigeons fly into m pigeonholes and $n > m$, then at least one hole must contain two or more pigeons. This principle is illustrated in Figure 9.4.1 for $n = 5$ and $m = 4$.

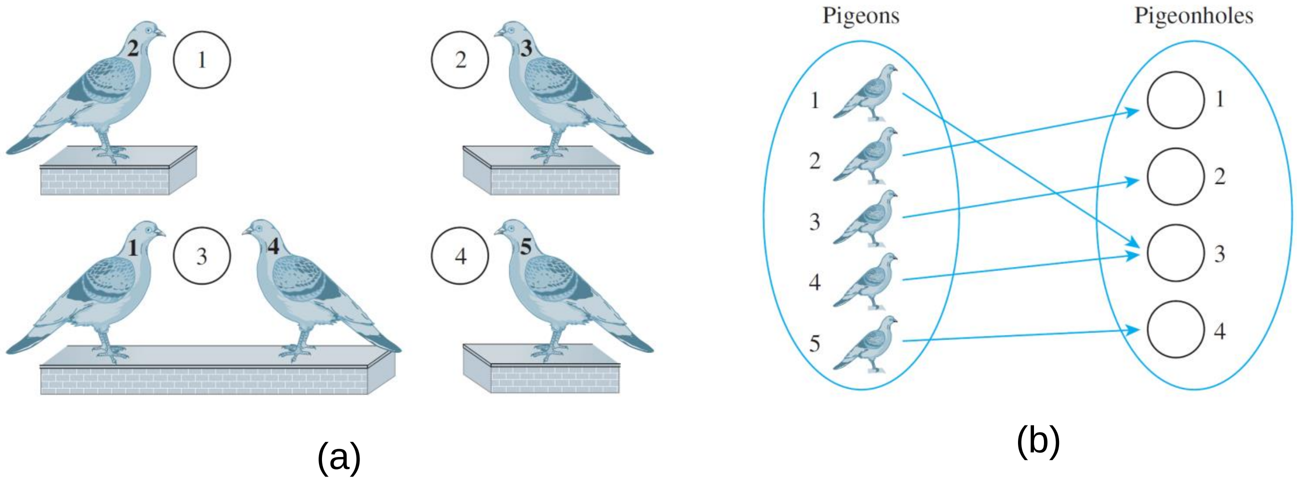


Figure 9.4.1

The Pigeonhole Principle

Illustration (a) shows the pigeons perched next to their holes, and (b) shows the correspondence from pigeons to pigeonholes.

The pigeonhole principle is sometimes called the ***Dirichlet box principle*** because it was first stated formally by J. P. G. L. Dirichlet (1805–1859).

Illustration (b) suggests the following mathematical way to phrase the principle.

Pigeonhole Principle

A function from one finite set to a smaller finite set cannot be one-to-one: There must be a least two elements in the domain that have the same image in the co-domain.

Example 1 – *Applying the Pigeonhole Principle*

a. In a group of six people, must there be at least two who were born in the same month?
In a group of thirteen people, must there be at least two who were born in the same month? Why?

b. Among the residents of New York City, must there be at least two people with the same number of hairs on their heads? Why?

Example 1(a) – *Solution*

A group of six people need not contain two who were born in the same month. For instance, the six people could have birthdays in each of the six months January through June.

A group of thirteen people, however, must contain at least two who were born in the same month, for there are only twelve months in a year and 13 > 12. Think of the thirteen people as the pigeons, and the twelve months of the year as the pigeonholes.

Example 1(b) – *Solution*

The answer is yes.

In this example the pigeons are the people of New York City and the pigeonholes are all possible numbers of hairs on any individual's head.

Call the population of New York City P. It is known that P is at least 5,000,000.

Also the maximum number of hairs on any person's head is known to be no more than 300,000.

Example 1(b) – *Solution*

Since the number of people in New York City is larger than the number of possible hairs on their heads, the function H is not one-to-one; at least two arrows point to the same number.

But that means that at least two people have the same number of hairs on their heads.

Generalized Pigeonhole Principle

A generalization of the pigeonhole principle states that if n pigeons fly into m pigeonholes and, for some positive integer k, $k < n/m$, then at least one pigeonhole contains $k + 1$ or more pigeons.

Generalized Pigeonhole Principle

For any function f from a finite set X with n elements to a finite set Y with m elements and for any positive integer k, if $k < n/m$, then there is some $y \in Y$ such that y is the image of at least $k + 1$ distinct elements of X.

Example 5 – *Applying the Generalized Pigeonhole Principle*

Show how the generalized pigeonhole principle implies that in a group of 85 people, at least 4 must have the same last initial.

Solution:

In this example the pigeons are the 85 people and the pigeonholes are the 26 possible last initials of their names. Note that $3 < 85/26 \cong 3.27$. Since $3 < 85/26$, the generalized pigeonhole principle states that some initial must be the image of at least four $(3 + 1)$ people. Thus, at least four people have the same last initial.

In-class Assignment #2

- If (n + 1) integers are chosen from the set
  {1, 2, 3, …, 2n},
  where n is a positive integer, must at least one of them be **odd?** Why?
  Must at least one of them be **even?** Why?