Chapter 7. Inclusion-Exclusion
a.k.a. The Sieve Formula

Prof. Tesler

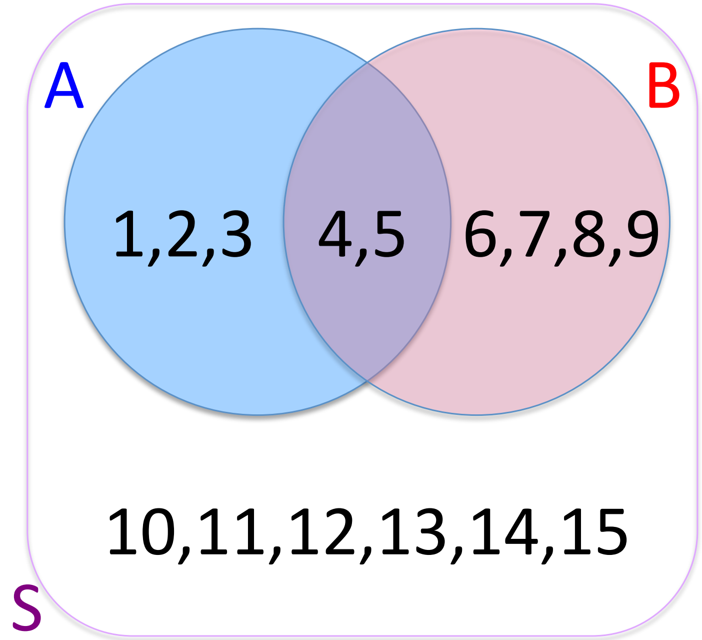
Math 184A
Fall 2019

Venn diagram and set sizes

$$A = \{1, 2, 3, 4, 5\}$$
$$B = \{4, 5, 6, 7, 8, 9\}$$

$$A \cup B = \{1, \ldots, 9\}$$
$$A \cap B = \{4, 5\}$$

$$(A \cup B)^c = \{10, \ldots, 15\}$$



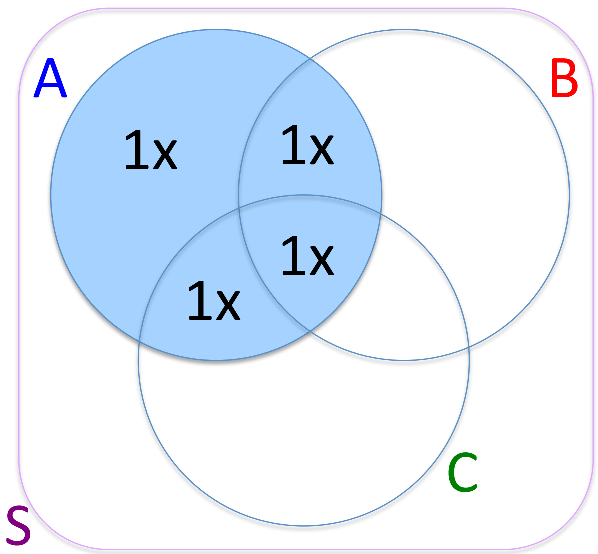
- $|A| + |B|$ counts everything in the union, but elements in the intersection are counted twice. Subtract $|A \cap B|$ to compensate:

$$\begin{array}{ccccccc} |A \cup B| & = & |A| & + & |B| & - & |A \cap B| \\ 9 & = & 5 & + & 6 & - & 2 \end{array}$$

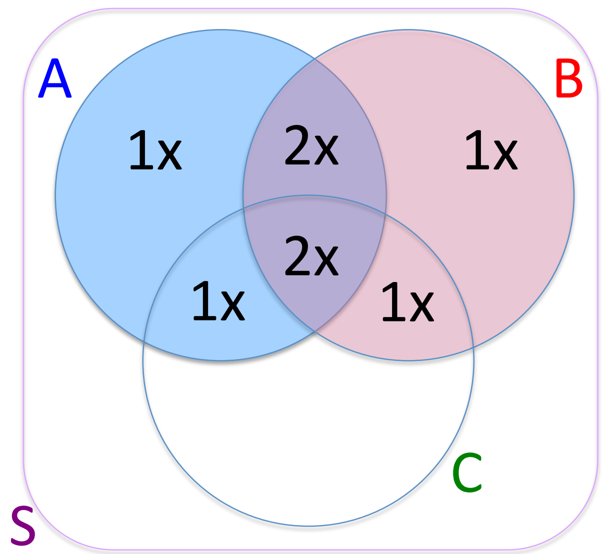
- Size of outside region:

$$\begin{array}{ccccccccccccc} |(A \cup B)^c| & = & |S| & - & |A \cup B| & = & |S| & - & |A| & - & |B| & + & |A \cap B| \\ 6 & = & 15 & - & 9 & = & 15 & - & 5 & - & 6 & + & 2 \end{array}$$

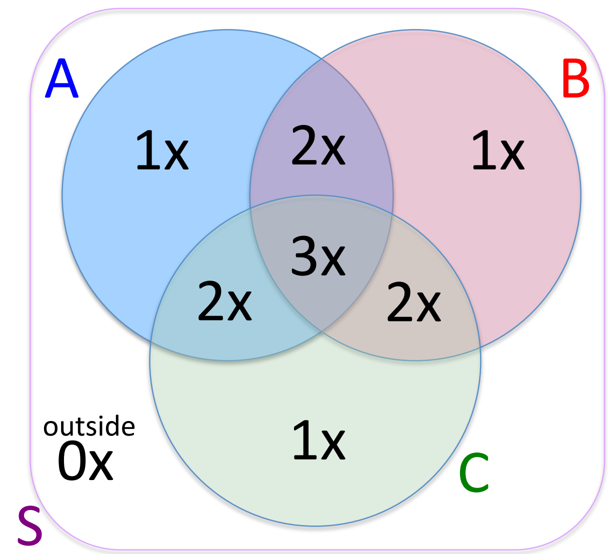
Size of a 3-way union



$|A|$



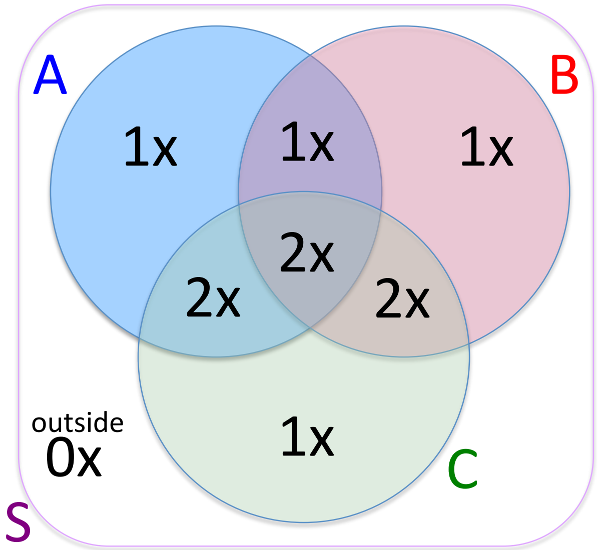
$|A| + |B|$



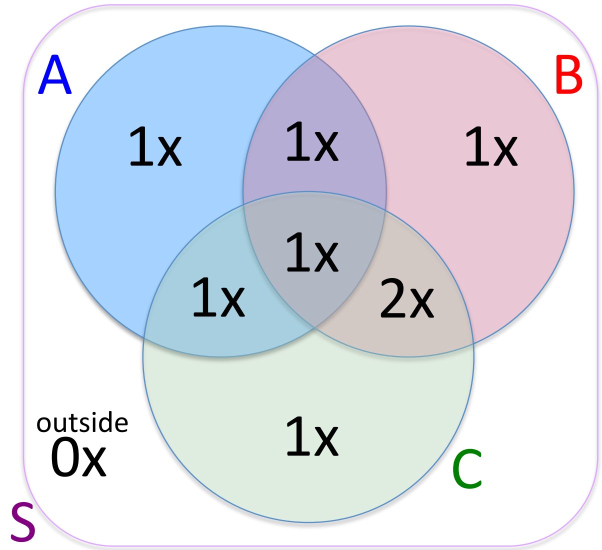
$|A| + |B| + |C|$

2x means the region is counted times.

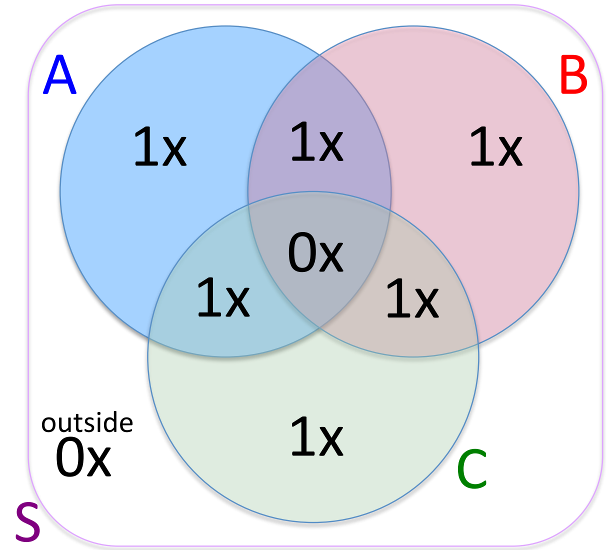
Size of a 3-way union



$$|A| + |B| + |C|$$
$$- |A \cap B|$$

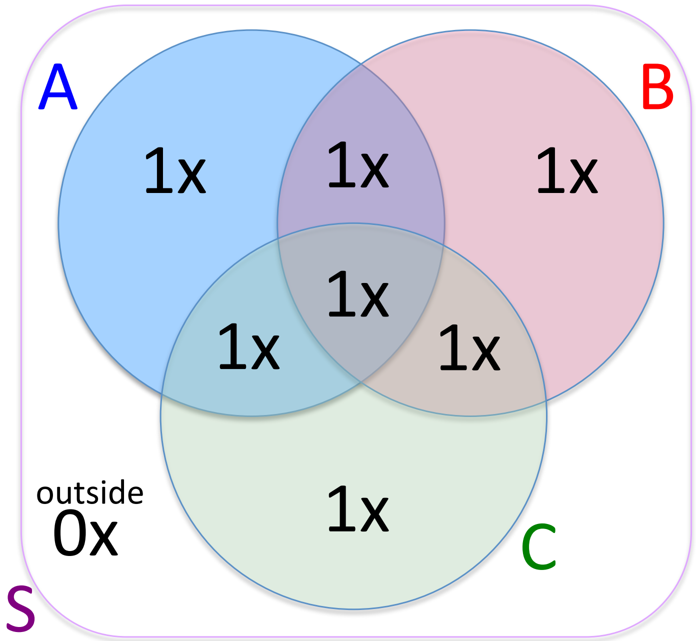


$$|A| + |B| + |C|$$
$$- |A \cap B|$$
$$- |A \cap C|$$



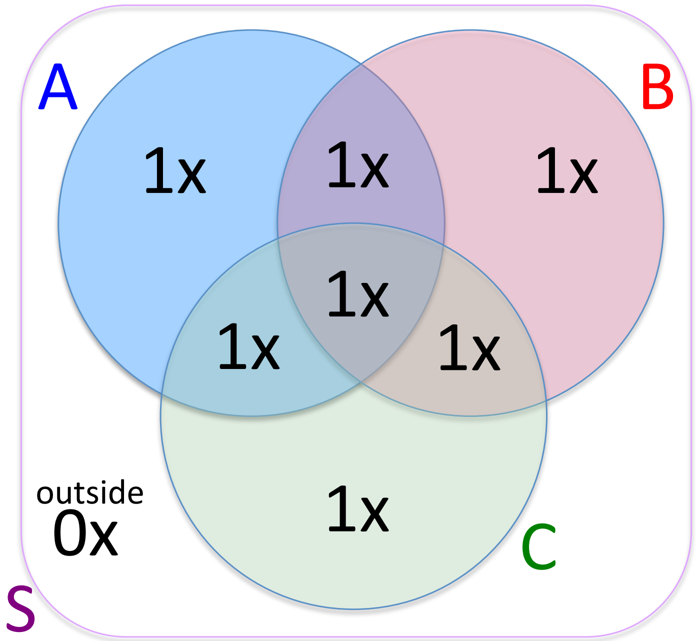
$$|A| + |B| + |C|$$
$$- |A \cap B|$$
$$- |A \cap C|$$
$$- |B \cap C|$$

Size of a 3-way union



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

Size of a 3-way union



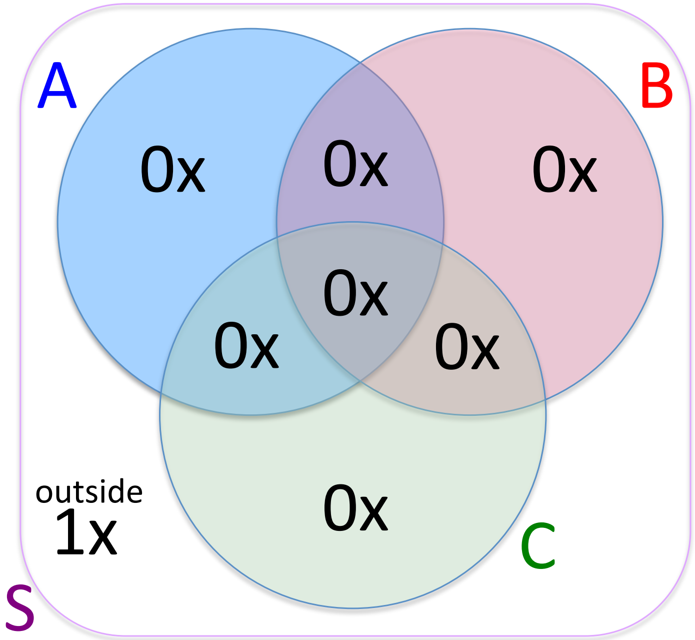
$$
\begin{aligned}
|A \cup B \cup C| &= (|A| + |B| + |C|) \\
&\quad - (|A \cap B| + |A \cap C| + |B \cap C|) \\
&\quad + |A \cap B \cap C| \\
&= N_1 - N_2 + N_3
\end{aligned}
$$

where N_i is the sum of sizes of i-way intersections:

$$
\begin{aligned}
N_1 &= |A| + |B| + |C| \\
N_2 &= |A \cap B| + |A \cap C| + |B \cap C| \\
N_3 &= |A \cap B \cap C|
\end{aligned}
$$

This is called inclusion-exclusion since we alternately include some parts, then exclude parts, then include parts, . . .

Size of complement of a 3-way union



$$
\begin{aligned}
|(A \cup B \cup C)^c| &= |S| - |A \cup B \cup C| \\
&= |S| \\
&\quad - (|A| + |B| + |C|) \\
&\quad + (|A \cap B| + |A \cap C| + |B \cap C|) \\
&\quad - |A \cap B \cap C| \\
&= N_0 - N_1 + N_2 - N_3
\end{aligned}
$$

where $N_0 = |S|$.

Inclusion-Exclusion Formula for size of union of n sets

a.k.a. The Sieve Formula

Inclusion-Exclusion Theorem:

- Let $A_1, \ldots, A_n$ be subsets of a finite set S.
- Let $N_0 = |S|$ and N_j be the sum of sizes of all j-way intersections

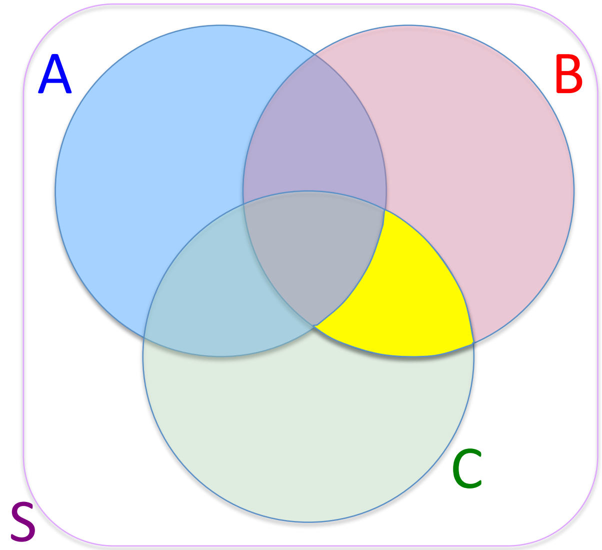
$$N_j = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_j \leqslant n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}| \qquad \text{for } j = 1, \ldots, n$$

- Then

$$|A_1 \cup \cdots \cup A_n| = N_1 - N_2 + N_3 - N_4 \cdots \pm N_n = \sum_{j=1}^{n} (-1)^{j-1} N_j$$

$$|(A_1 \cup \cdots \cup A_n)^c| = N_0 - N_1 + N_2 - N_3 + N_4 \cdots \mp N_n = \sum_{j=0}^{n} (-1)^{j} N_j$$

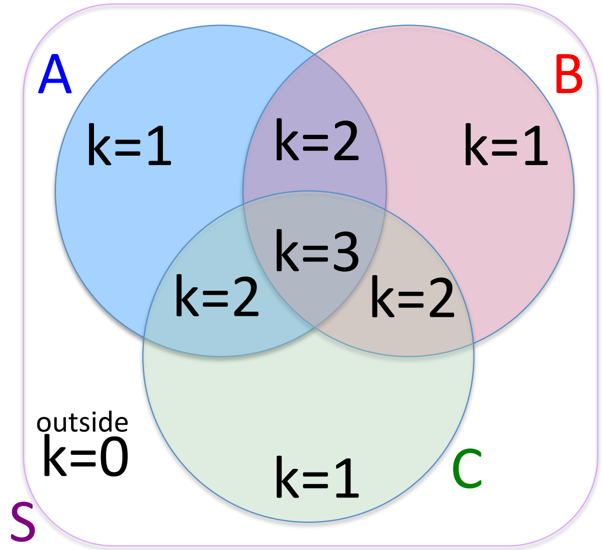
Proof of Inclusion-Exclusion Formula



- The yellow region is inside $k = 2$ sets (B and C)
  and outside $n - k = 3 - 2 = 1$ set (A).
- Which j-way intersections among A,B,C contain the yellow region? The ones that only involve B and/or C. None that involve A.
- For each j, it's in $\binom{k}{j}$ j-way intersections:

$$j = 1: \quad \binom{2}{1} = 2 \quad B \text{ alone and } C \text{ alone}$$
$$j = 2: \quad \binom{2}{2} = 1 \quad B \cap C$$
$$j = 3: \quad \binom{2}{3} = 0 \quad \text{None}$$

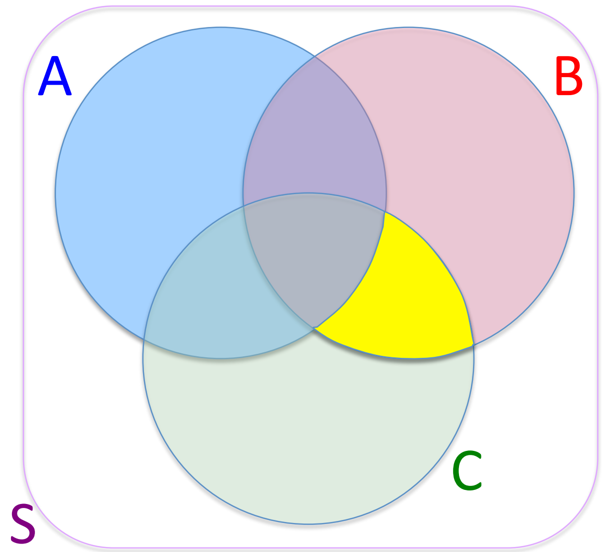
Proof of Inclusion-Exclusion Formula



A region in exactly k of the sets $A_1, \ldots, A_n$ is counted in $N_1 - N_2 + N_3 - \cdots$ this many times (shown for $n = 3$):

	N_1	$-$	N_2	$+$	N_3		
Contribution	$\binom{k}{1}$	$-$	$\binom{k}{2}$	$+$	$\binom{k}{3}$	$=$	
$k=0:$	$\binom{0}{1}$	$-$	$\binom{0}{2}$	$+$	$\binom{0}{3}$	$=$	$0-0+0=0$
$k=1:$	$\binom{1}{1}$	$-$	$\binom{1}{2}$	$+$	$\binom{1}{3}$	$=$	$1-0+0=1$
$k=2:$	$\binom{2}{1}$	$-$	$\binom{2}{2}$	$+$	$\binom{2}{3}$	$=$	$2-1+0=1$
$k=3:$	$\binom{3}{1}$	$-$	$\binom{3}{2}$	$+$	$\binom{3}{3}$	$=$	$3-3+1=1$

Proof of Inclusion-Exclusion Formula



- In general, consider a region R of the Venn diagram
  inside k of the sets $A_1, \ldots, A_n$ (call them $I_1, \ldots, I_k$)
  and outside the other $n-k$ (call them $O_1, \ldots, O_{n-k}$).
- The j-way intersections of A's that R is in use any j of the I's and none of the O's. Thus, R is in $\binom{k}{j}$ j-way intersections.
- All elements of R are counted

  $\binom{k}{j}$ times in N_j

  and $\sum_{j=1}^{n}(-1)^{j-1}\binom{k}{j}$ times in $\sum_{j=1}^{n}(-1)^{j-1}N_j$.

Pascal's Triangle

- This is related to alternating sums in Pascal's Triangle:

For $n = 3$:

	Contribution	**Alternating sum in Pascal's Triangle**
$k = 0:$	$0 - 0 + 0 = 0$	$1 \qquad = 1$
$k = 1:$	$1 - 0 + 0 = 1$	$1 - 1 \qquad = 0$
$k = 2:$	$2 - 1 + 0 = 1$	$1 - 2 + 1 \qquad = 0$
$k = 3:$	$3 - 3 + 1 = 1$	$1 - 3 + 3 - 1 = 0$

For $n \geqslant 0$:

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

$$\sum_{j=0}^{k} (-1)^{j} \binom{k}{j} = \begin{cases} 1 & \text{if } k = 0; \\ 0 & \text{if } k > 0. \end{cases}$$

- The summation limits differ and the signs are opposite. (Also, the variable names differ from earlier slides on Pascal's Triangle.)

Proof of Inclusion-Exclusion Formula

- A Venn diagram region R inside k of the sets $A_1, \ldots, A_n$ and outside the other $n-k$ is counted $\sum_{j=1}^{n} (-1)^{j-1} \binom{k}{j}$ times in $\sum_{j=1}^{n} (-1)^{j-1} N_j$.

- This multiplicity is related to $(1-1)^k = \sum_{j=0}^{k} (-1)^j \binom{k}{j}$.
  - Since $\binom{k}{j} = 0$ for $j > k$, we can extend the sum up to $j = n$:
  $$(1-1)^k = \sum_{j=0}^{n} (-1)^j \binom{k}{j}$$
  - The $j = 0$ term is $(-1)^0 \binom{k}{0} = 1$. Subtract from 1:
  $$1 - (1-1)^k = -\sum_{j=1}^{n} (-1)^j \binom{k}{j} = \sum_{j=1}^{n} (-1)^{j-1} \binom{k}{j}$$

- For $k > 0$, the # times R is counted is $1 - (1-1)^k = 1 - 0^k = 1$.

- For $k = 0$ (outside region): All $\binom{k}{j} = 0$, so $\sum_{j=1}^{n} (-1)^{j-1} \binom{k}{j} = 0$.

- Thus, all regions of $A_1 \cup \cdots \cup A_n$ are counted once, and the outside is not counted. QED.

Derangements

A class with n students takes a pop quiz. Everyone has to give their test to someone else to grade. Each person just gets one test to grade, and it can't be their own.

How many ways are there to do this? Call it D_n.

Derangements

- A *fixed point* of a function $f(x)$ is a point where $f(x) = x$.
- One-line notation for a permutation: 24135 represents
  $f(1) = 2 \qquad f(2) = 4 \qquad f(3) = 1 \qquad f(4) = 3 \qquad f(5) = 5$
- 5 is a fixed point since $f(5) = 5$.
- A *derangement* is a permutation with no fixed points.
- Let D_n be the number of derangements of size n.

Derangements: Examples

- $n = 1$: There are none! The only permutation is $f(1) = 1$, which has a fixed point. So $D_1 = 0$.
- $n = 2$: 21, so $D_2 = 1$.
- $n = 3$: 231, 312, so $D_3 = 2$.
- $n = 4$: 2143, 2341, 2413, 3142, 3412, 3421, 4123, 4312, 4321, so $D_4 = 9$.
- $n = 0$: This is a vacuous case. The empty function $f : \emptyset \to \emptyset$ does not have any fixed points, so $D_0 = 1$.

Derangements: Formula with Inclusion-Exclusion

- Let S be the set of all permutations on $[n]$.
- For $i = 1, \ldots, n$, let $A_i \subseteq S$ be all permutations with $f(i) = i$.
  The other $n-1$ elements can be permuted arbitrarily, so $|A_i| = (n-1)!$.
- The set of all derangements of $[n]$ is $(A_1 \cup \cdots \cup A_n)^c$.
- We will use Inclusion-Exclusion to compute the size of this as

$$|(A_1 \cup \cdots \cup A_n)^c| = \sum_{j=0}^{n} (-1)^j N_j$$

  We need to compute the N_j's for this.

Derangements: Formula with Inclusion-Exclusion

$S =$ set of all permutations on $[n]$

For $i = 1, \ldots, n$: $A_i =$ set of permutations with $f(i) = i$

Consider the 3-way intersection $A_1 \cap A_4 \cap A_8$:

- It consists of permutations with $f(1) = 1$, $f(4) = 4$, $f(8) = 8$, and any permutation of $[n] \setminus \{1, 4, 8\}$ in the remaining entries.
- The number of such permutations is $(n-3)!$.

In general:

- Every 3-way intersection $A_{i_1} \cap A_{i_2} \cap A_{i_3}$ has size $(n-3)!$.
- There are $\binom{n}{3}$ three-way intersections. The sum of their sizes is

$$N_3 = \binom{n}{3} \cdot (n-3)! = \frac{n!}{3!\,(n-3)!} \cdot (n-3)! = \frac{n!}{3!}$$

- For $j = 1, \ldots, n$, we similarly get $N_j = n!/j!$.
- Also, $N_0 = |S| = n! = n!/0!$. Thus,

$$D_n = |(A_1 \cup \cdots \cup A_n)^c| = \sum_{j=0}^{n} (-1)^j \, \frac{n!}{j!}$$

Derangements: Formula with Inclusion-Exclusion

Derangements for $n = 4$:

2143, 2341, 2413, 3142, 3412, 3421, 4123, 4312, 4321

$$
\begin{aligned}
D_4 &= \frac{4!}{0!} - \frac{4!}{1!} + \frac{4!}{2!} - \frac{4!}{3!} + \frac{4!}{4!} \\
&= 24 - 24 + 12 - 4 \quad + 1 \quad = 9
\end{aligned}
$$

Second formula for D_n

- Factor out $n!$:

$$D_n = \sum_{j=0}^{n} (-1)^j \frac{n!}{j!} = n! \sum_{j=0}^{n} \frac{(-1)^j}{j!}$$

- This resembles

$$e^{-1} = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!}$$

- So

$$\frac{n!}{e} = \sum_{j=0}^{\infty} (-1)^j \frac{n!}{j!} = \underbrace{\sum_{j=0}^{n} (-1)^j \frac{n!}{j!}}_{=D_n} + \underbrace{\sum_{j=n+1}^{\infty} (-1)^j \frac{n!}{j!}}_{\text{call this } Q}$$

Second formula for D_n

$$\frac{n!}{e} = \sum_{j=0}^{\infty} (-1)^j \frac{n!}{j!} = \underbrace{\sum_{j=0}^{n} (-1)^j \frac{n!}{j!}}_{=D_n} + \underbrace{\sum_{j=n+1}^{\infty} (-1)^j \frac{n!}{j!}}_{\text{call this } Q}$$

- Terms in Q alternate in sign and strictly decrease in magnitude, so

$$|Q| < |\text{first term}| = \frac{n!}{(n+1)!} = \frac{1}{n+1}$$

For $n \geqslant 1$, this gives $|Q| < 1/2$. So D_n is within $\pm 1/2$ of $n!/e$.

Theorem

- **Theorem:** For $n \geqslant 1$, D_n is $\frac{n!}{e}$ rounded to the closest integer.
  For $n = 0$, use $D_0 = 1$.
- **Example:** For $n = 4$, compute $4!/e \approx 8.829$, which rounds to $D_4 = 9$.
- For $n = 0$, rounding doesn't work: $\text{round}(0!/e) = \text{round}(0.3678794) = 0$, but $D_0 = 1$.

Recursion for D_n

- Observe

$$D_4 = \tfrac{4!}{0!} - \tfrac{4!}{1!} + \tfrac{4!}{2!} - \tfrac{4!}{3!} + \tfrac{4!}{4!} = 4\left(\tfrac{3!}{0!} - \tfrac{3!}{1!} + \tfrac{3!}{2!} - \tfrac{3!}{3!}\right) + \tfrac{4!}{4!} = 4D_3 + 1$$

- For $n \geqslant 1$:

$$D_n = \sum_{j=0}^{n} (-1)^j \, \frac{n!}{j!} = \left(\sum_{j=0}^{n-1} (-1)^j \, \frac{n!}{j!} \right) + (-1)^n \, \frac{n!}{n!}$$

$$= n \left(\sum_{j=0}^{n-1} (-1)^j \, \frac{(n-1)!}{j!} \right) + (-1)^n = n \, D_{n-1} + (-1)^n$$

- Use initial condition $D_0 = 1$
  and recursion $D_n = n \, D_{n-1} + (-1)^n$ for $n \geqslant 1$:

$$\begin{aligned}
D_0 &= 1 \\
D_1 &= 1D_0 - 1 = 1(1) - 1 = 0 \\
D_2 &= 2D_1 + 1 = 2(0) + 1 = 1 \\
D_3 &= 3D_2 - 1 = 3(1) - 1 = 2 \\
D_4 &= 4D_3 + 1 = 4(2) + 1 = 9
\end{aligned}$$

How many surjections $f : [n] \to [k]$?

Recall from Chapter 5 that the number of surjections $f : [n] \to [k]$ is $k!\,S(n,k)$. Here is a different formula, using inclusion-exclusion.

- Let S be the set of all functions $f : [n] \to [k]$. So $|S| = k^n$.
- For $i = 1, \ldots, k$, let A_i be the set of all functions with $f^{-1}(i) = \emptyset$.
- The set of surjections $f : [n] \to [k]$ is $(A_1 \cup \cdots \cup A_k)^c$.
- A_i has $k-1$ choices of how to map each of $f(1), \ldots, f(n)$, so
$$|A_i| = (k-1)^n.$$
- $A_1 \cap A_2 \cap A_3$ is the set of functions with nothing mapped to 1, 2, or 3. There are $k-3$ choices of how to map each of $f(1), \ldots, f(n)$, so
$$|A_1 \cap A_2 \cap A_3| = (k-3)^n.$$
- The sum of all sizes of all 3-way intersections of $A_1 \ldots, A_k$ is
$$N_3 = \binom{k}{3}(k-3)^n.$$
- Similarly, for $j = 0, \ldots, k$,
$$N_j = \binom{k}{j}(k-j)^n.$$

How many surjections $f : [n] \to [k]$?

$S =$ set of all functions $f : [n] \to [k]$

For $i = 1, \ldots, k$: $A_i =$ functions with $f^{-1}(i) = \emptyset$

Size of j-way intersections: $N_j = \binom{k}{j}(k-j)^n$

- Thus, the number of surjections is

$$|(A_1 \cup \cdots \cup A_k)^c| = \sum_{j=0}^{k} (-1)^j N_j = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n$$

Example

- For $n = 3$ and $k = 2$,

$$\sum_{j=0}^{2} (-1)^j \tbinom{2}{j} (2-j)^3 = \tbinom{2}{0}(2-0)^3 - \tbinom{2}{1}(2-1)^3 + \tbinom{2}{2}(2-2)^3$$
$$= 8 - 2 + 0 = 6$$

- So there are 6 surjections $f : [3] \to [2]$.
- In one-line notation, they are: 112, 121, 211, 122, 212, 221

Formula for $S(n,k)$

Corollary

Since the number of surjections also equals $k!\,S(n,k)$*:*

$$S(n,k) = \frac{1}{k!}\sum_{j=0}^{k}(-1)^j \binom{k}{j}(k-j)^n$$

Continuing the example, $S(3,2) = 6/2! = 3$.