COMP170
Discrete Mathematical Tools for Computer Science

Inclusion-Exclusion

Version 1.2: Last updated, November 20, 2006

Discrete Math for Computer Science
K. Bogart, C. Stein and R.L. Drysdale
Section 5.2, pp. 224-233

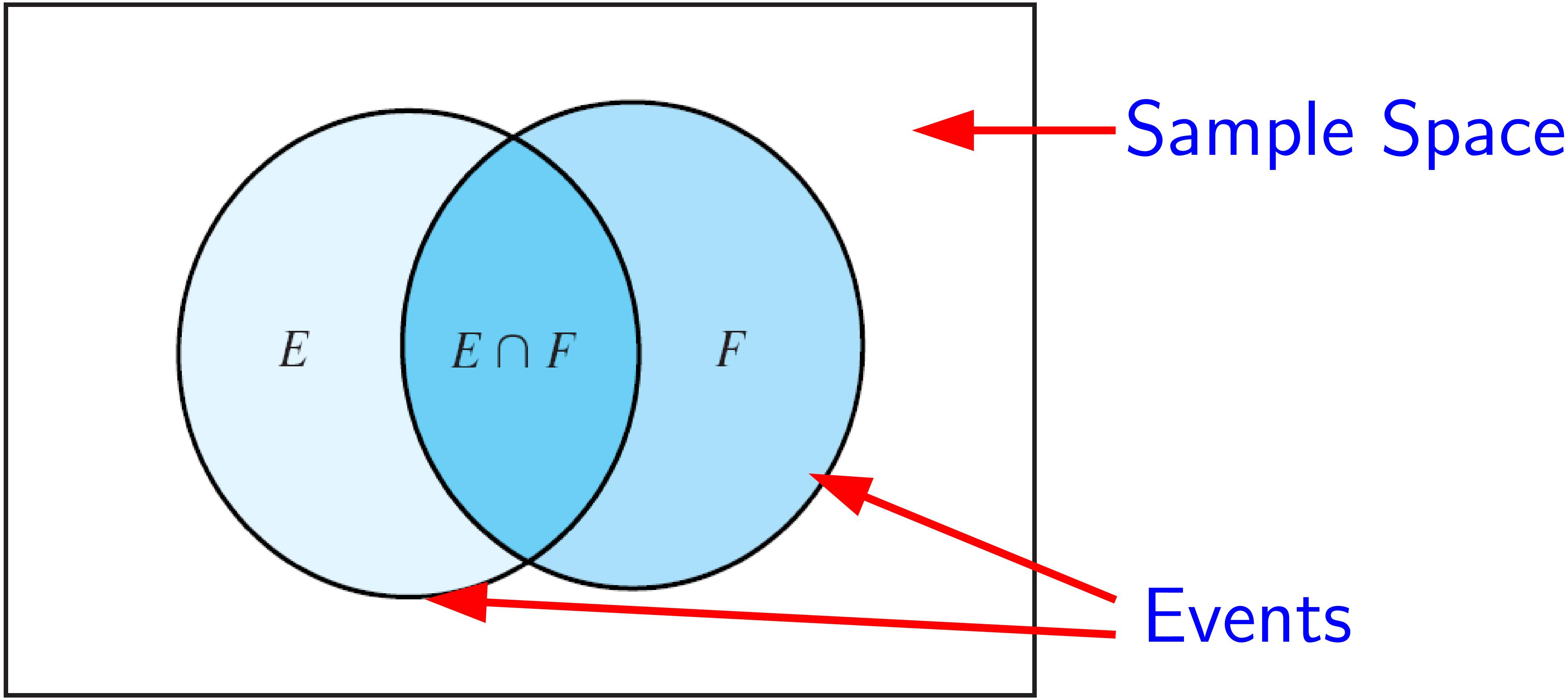
Unions and Intersections

- The Probability of a Union of Events
- The Principle of Inclusion and Exclusion for Probability
- The Principle of Inclusion and Exclusion for Counting

The Probability of a Union of Events

In $P(E) + P(F)$, weights of elements of $E \cap F$ each appear **twice**, while weights of all other elements of $E \cup F$ each appear exactly **once**.

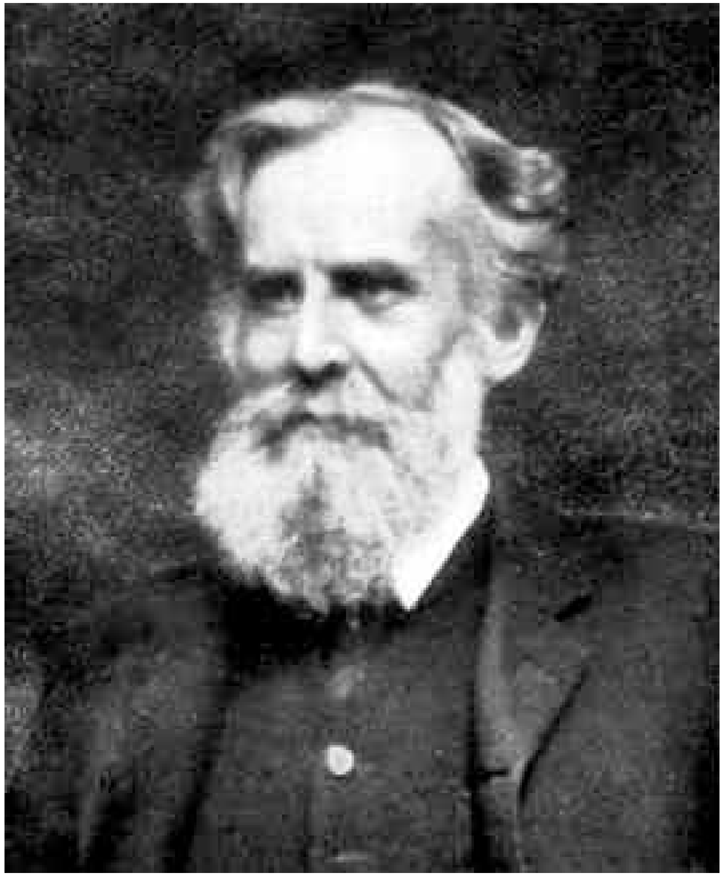
Venn Diagram



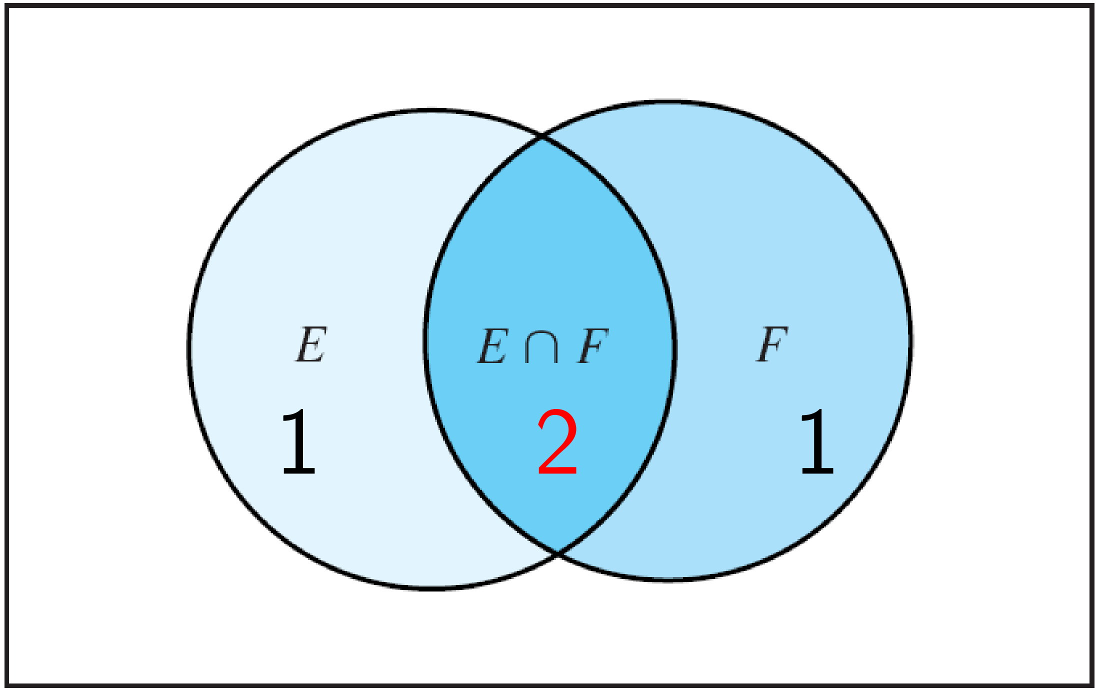
John Venn

b. 1834, d. 1923

British Mathematician who continued the work of Boole. Although he was not the first person to use diagrams in formal logic, he seems to have been the first to formalize their usage and generalize them.



For more, see the survey of Venn diagrams at http://www.combinatorics.org/Surveys/ds5/VennJohnEJC.html



$P(E) + P(F)$ counts probability weights of each element of $E \cap F$ **twice**.

Thus, to get a sum that includes probability weight of each element of $E \cup F$ exactly **once**, we must **subtract** weight of $E \cap F$ from $P(E) + P(F)$.

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) \qquad \textbf{(*)}$$

If you roll two dice, what is the probability of either an even sum or a sum of 8 or more (or both)?

Event E: Sum is even

Event F: Sum is 8 or more

$$P(E) = \frac{1}{2}$$

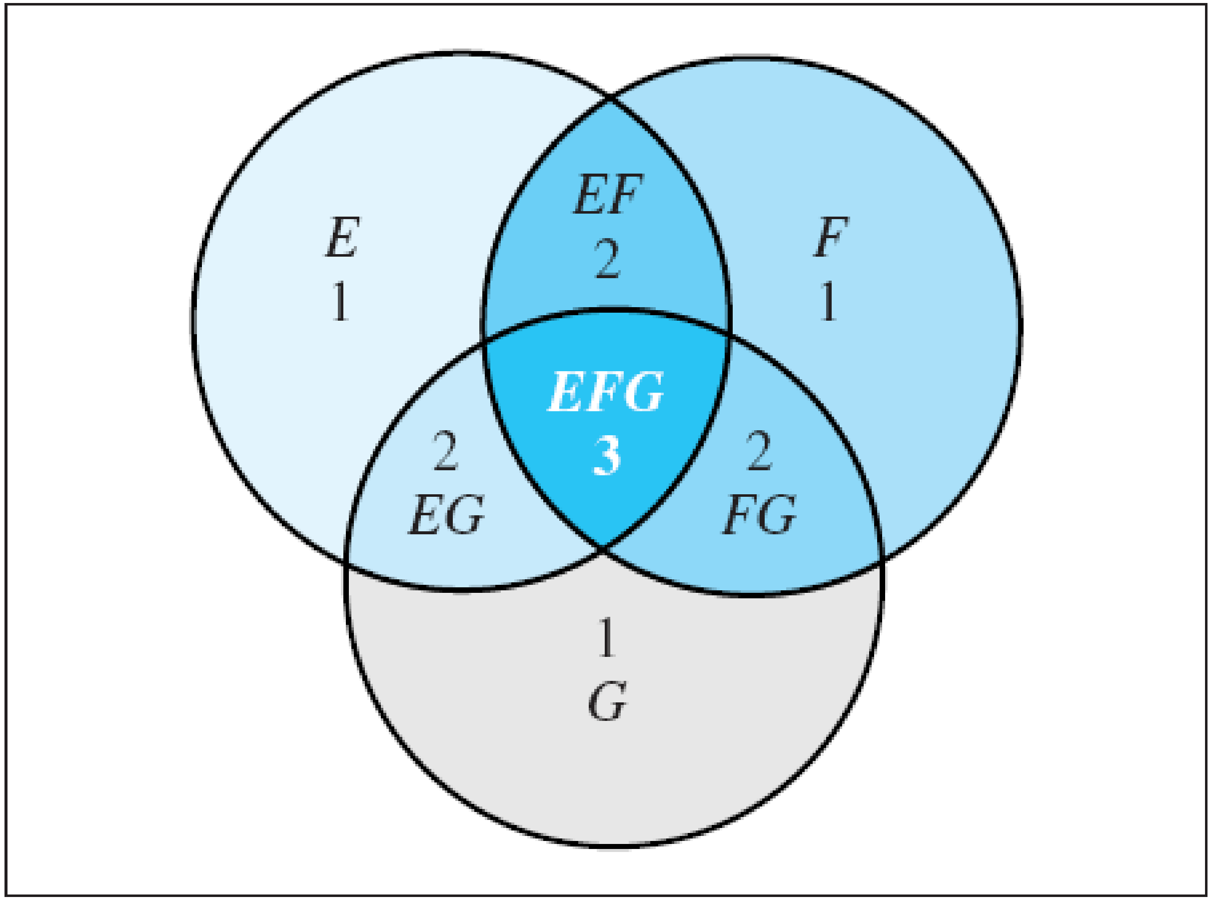
$$P(F) = \underset{P(8)}{\frac{5}{36}} + \underset{P(9)}{\frac{4}{36}} + \underset{P(10)}{\frac{3}{36}} + \underset{P(11)}{\frac{2}{36}} + \underset{P(12)}{\frac{1}{36}} = \frac{15}{36}$$

Probability of **even** sum of 8 or more is

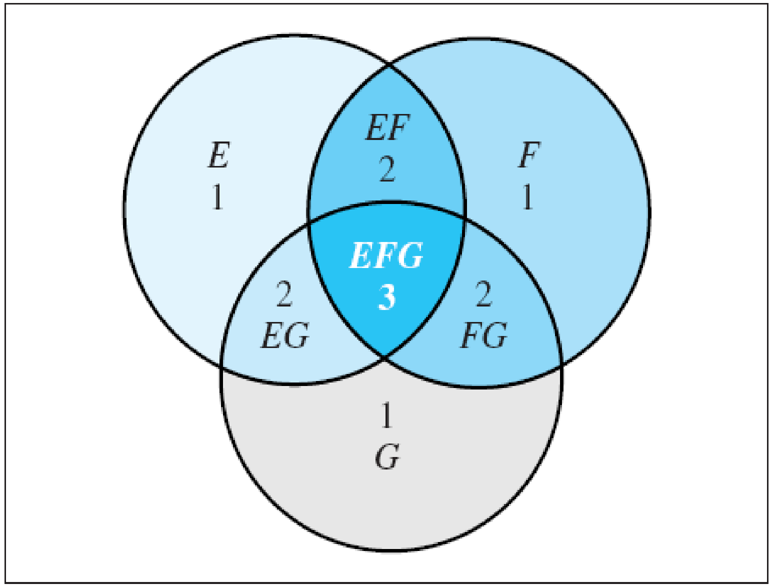
$$P(E \cap F) = \underset{P(8)}{\frac{5}{36}} + \underset{P(10)}{\frac{3}{36}} + \underset{P(12)}{\frac{1}{36}} = \frac{9}{36}.$$

$$\Rightarrow \; P(E \cup F) = \boxed{P(E) + P(F) - P(E \cap F) = \frac{1}{2} + \frac{15}{36} - \frac{9}{36} = \frac{2}{3}}$$

The Union of Three events: $E \cup F \cup G$



When adding $P(E) + P(F) + P(G)$,
weights of elements in regions $E \cap F$, $F \cap G$, and $E \cap G$
but not $E \cap F \cap G$, are counted exactly twice but
weights of elements in $E \cap F \cap G$,
are counted exactly three times



Want to calculate $P(E \cup F \cup G)$.

Start with $P(E) + P(F) + P(G)$.

This
Double counts events in EF, EG, FG
Triple counts events in EFG

Subtracting weights of elements of each $E \cap F$, $F \cap G$, and $E \cap G$ doesn't quite work, since this subtracts weights of elements in EF, FG, and EG once (good) but also subtracts weights of elements in EFG three times (bad).

So, add weights of elements in $E \cap F \cap G$ back into our sum.

$$\boxed{\begin{aligned} P(E \cup F \cup G) = {} & P(E) + P(F) + P(G) \\ & - P(E \cap F) - P(E \cap G) - P(F \cap G) \\ & + P(E \cap F \cap G). \end{aligned}}$$

Unions and Intersections

- The Probability of a Union of Events
- The Principle of Inclusion and Exclusion for Probability
- The Principle of Inclusion and Exclusion for Counting

Principle of Inclusion and Exclusion for Probability

So far we've seen:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(E \cap F) - P(E \cap G) - P(F \cap G) + P(E \cap F \cap G)$$

We now guess the general formula:

$$\boxed{P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(E_i \cap E_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(E_i \cap E_j \cap E_k) - \ldots .}$$

Our guess:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} P(E_i \cap E_j) + \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} P(E_i \cap E_j \cap E_k) - \ldots.$$

To prove this we'll introduce new notation for sums:

Denote sum over all increasing sequences $i_1, i_2, \ldots, i_k$ of integers between 1 and n, of probs of sets $E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k}$ by:

$$\boxed{\sum_{\boxed{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \ldots < i_k \le n}}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})}$$

More generally:

$$\sum_{\substack{i_1,i_2,\ldots,i_k:\\ 1\le i_1<i_2<\ldots<i_k\le n}} f(i_1, i_2, \ldots, i_k)$$

is the sum of $f(i_1, i_2, \ldots, i_k)$ over all increasing sequences of k numbers between 1 and n.

Example:

What is

$$\sum_{\substack{i_1,i_2,i_3:\\ 1\le i_1<i_2<i_3\le 4}} \left(i_1 + i_2 + i_3\right) \quad ?$$

$$(1+2+3)+(1+2+4)+(1+3+4)+(2+3+4)$$
$$= 6+7+8+9 = 30.$$

Theorem 5.3
(Principle of Inclusion and Exclusion for Probability)
The probability of the union $E_1 \cup E_2 \cup \ldots \cup E_n$ of events in a sample space S is given by

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k : \\ 1 \le i_1 < i_2 < \ldots < i_k \le n}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$$

Note: we've already seen $n = 2, 3$:

$n = 2$ $\quad P(E \cup F) = P(E) + P(F) - P(E \cap F)$

$n = 3$ $\quad P(E \cup F \cup G) = P(E) + P(F) + P(G)$
$$- P(E \cap F) - P(E \cap G) - P(F \cap G)$$
$$+ P(E \cap F \cap G)$$

Theorem 5.3
(Principle of Inclusion and Exclusion for Probability)
The probability of the union $E_1 \cup E_2 \cup \ldots \cup E_n$
of events in a sample space S is given by

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \ldots < i_k \le n}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$$

Proof (by mathematical induction):

Note: Book also gives a 2^{nd}, combinatorial proof

Base case $n = 2$:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Suspose inductively that for any family of $n-1$ sets $F_1, F_2, \ldots, F_{n-1}$,

$$P\left(\bigcup_{i=1}^{n-1} F_i\right) = \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \ldots < i_k \le n-1}} P(F_{i_1} \cap F_{i_2} \cap \ldots \cap F_{i_k})$$

Now assume we have family $E_1, E_2, \ldots, E_n$ of n sets.

Set $E = E_1 \cup \ldots \cup E_{n-1}$ and $F = E_n$.
Then, by $P(E \cup F) = P(E) + P(F) - P(E \cap F)$ (i.h., $n = 2$)

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P(E_n) - P\left(\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n\right)$$

So Far

$E = E_1 \cup \ldots \cup E_{n-1}$ and $F = E_n$.

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P(E_n) - \boxed{P\left(\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n\right)}$$

First term on RHS is given by i.h.

To get 3^{rd} term we note, from distributive law, that

$$\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n = \bigcup_{i=1}^{n-1} (E_i \cap E_n)$$

So Far

$E = E_1 \cup \ldots \cup E_{n-1}$ and $F = E_n$.

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P(E_n) - \boxed{P\left(\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n\right)}$$

First term on RHS is given by i.h.

To get 3^{rd} term we note, from distributive law, that

Now, for $i < n$, set $G_i = E_i \cap E_n$. This gives

$$\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n = \bigcup_{i=1}^{n-1} (E_i \cap E_n) = \bigcup_{i=1}^{n-1} G_i$$

So Far

$E = E_1 \cup \ldots \cup E_{n-1}$ and $F = E_n$.

For $i < n$, set $G_i = E_i \cap E_n$.

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P(E_n) - P\left(\bigcup_{i=1}^{n-1} G_i\right)$$

We can now use i.h. to evauate the last term on the RHS. To do this, we will need to note that (why?)

$$-(-1)^{k+1}\ P(G_{i_1} \cap G_{i_2} \cap \ldots \cap G_{i_k})$$

$$= (-1)^{k+2}\ P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k} \cap E_n)$$

So Far

$E = E_1 \cup \ldots \cup E_{n-1}$ and $F = E_n$.

For $i < n$, set $G_i = E_i \cap E_n$.

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P(E_n) - P\left(\bigcup_{i=1}^{n-1} G_i\right)$$

Applying i.h. once and again

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \ldots < i_k \le n-1}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$$

$$+P(E_n) \quad + \sum_{k=1}^{n-1} (-1)^{k+2} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \ldots < i_k \le n-1}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k} \cap E_n)$$

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n-1}(-1)^{k+1} \sum_{\substack{i_1,i_2,\ldots,i_k: \\ 1\le i_1<i_2<\ldots<i_k\le n-1}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$$

$$+P(E_n) \quad + \sum_{k=1}^{n-1}(-1)^{k+2} \sum_{\substack{i_1,i_2,\ldots,i_k: \\ 1\le i_1<i_2<\ldots<i_k\le n-1}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k} \cap E_n)$$

First summation on RHS sums $(-1)^{k+1}P(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_k})$ over all lists $i_1, i_2, \ldots, i_k$ that **do not** contain n.

$P(E_n)$ and second summation together sums $(-1)^{k+1}P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$ over all lists $i_1, i_2, \ldots, i_k$ that **do** contain n.

Therefore,

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \ldots < i_k \le n}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$$

Thus, by principle of MI, formula holds for all $n > 1$.

Example:

$$
\begin{aligned}
P(E_1 \cup E_2 \cup E_3 \cup E_4) = \; & P(E_1) + P(E_2) + P(E_3) + P(E_4) \\
& -P(E_1 \cap E_2) - P(E_1 \cap E_3) - P(E_1 \cap E_4) \\
& -P(E_2 \cap E_3) - P(E_2 \cap E_4) - P(E_3 \cap E_4) \\
& +P(E_1 \cap E_2 \cap E_3) + P(E_1 \cap E_2 \cap E_4) \\
& +P(E_1 \cap E_3 \cap E_4) + P(E_2 \cap E_3 \cap E_4) \\
& -P(E_1 \cap E_2 \cap E_3 \cap E_4)
\end{aligned}
$$

Example:

There are n students who have the same model and color of back-pack. They went to a class and hung their backpacks up on the wall. Someone came along and totally mixed up the backpacks so the students get back random backpacks.

What is the probability that

(i) Exactly k specified students get their OWN backpacks back?
(ii) At least one student gets his/her OWN backpack back?
(iii) No student gets his/her OWN backpack back?

These problems are essentially equivalent to taking a random permutation f of $[1..n]$ and asking questions about:
for which numbers x is $f(x) = x$?

Problem (iii) is sometimes known as the derangement problem

This problem is equivalent to taking a random permutation f of $[1..n]$ and asking: for which numbers x is $f(x) = x$.

The sample space is the set S_n of all permutations of $[1..n]$

Note (why?) that there are exactly $(n-k)!$ permutations f s.t for k given numbers, $x_1, x_2, \ldots, x_k$, $f(x_i) = x_i$

$$\begin{aligned}\Rightarrow\ & P(k \text{ given students get their own backpack back}) \\ &= P(\text{ for } k \text{ given numbers } x_1, x_2, \ldots, x_k,\quad f(x_i) = x_i) \\ &= (n-k)!/n!\ .\end{aligned}$$

For later use, set $D_{n,k} = \frac{(n-k)!}{n!}$.

Note

If E_i is event that person i gets correct backpack back.

$\Rightarrow$
$$P(E_i) = D_{n,1} = \frac{(n-1)!}{n!} = \frac{1}{n}$$

Also note (why?)

$$P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k})$$

$$= P(\text{Students } i_1, i_2, \ldots, i_k \text{ get their backpacks back})$$

$$= \frac{(n-k)!}{n!} = D_{n,k}$$

Example $n = 5$:

Probability that at least one person gets his or her own backpack is $P(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5)$.

Then, by **principle of inclusion and exclusion**, probability that at least one person gets his or her own backpack is

$$P(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5)$$

$$= \sum_{k=1}^{5} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \leq i_1 < i_2 < \ldots < i_k \leq 5}} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k}) \quad (*)$$

Recall: there are $\binom{5}{k}$ sets of k people chosen from 5 students.

That is, there are $\binom{5}{k}$ lists $i_1, i_2, \ldots, i_k$ with
$1 \leq i_1 < i_2 < \ldots < i_k \leq 5$.

So rewrite RHS of (*) as $$\sum_{k=1}^{5} (-1)^{k+1} \binom{5}{k} \frac{(5-k)!}{5!}.$$

$$P(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5)$$

$$= \sum_{k=1}^{5} (-1)^{k+1} \binom{5}{k} \frac{(5-k)!}{5!}$$

$$= \sum_{k=1}^{5} (-1)^{k+1} \frac{5!}{k!(5-k)!} \frac{(5-k)!}{5!} = \sum_{k=1}^{5} (-1)^{k+1} \frac{1}{k!}$$

Probability that at least one person gets his or her own backpack is then

$$\sum_{k=1}^{5}(-1)^{k+1}\frac{1}{k!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!}$$

Probability that nobody gets his or her own backpack is 1 minus probability that someone does, or

$$\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} \qquad = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!}$$

General case n:

Probability of at least one person getting his or her own backpack is

$$\sum_{i=1}^{n}(-1)^{i+1}\frac{1}{i!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + \frac{(-1)^{n-1}}{n!}.$$

Probability nobody gets his or her own backpack is 1 minus the probability above, or

$$\sum_{i=2}^{n}(-1)^{i}\frac{1}{i!} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{n}}{n!}.$$

Recall from calculus:

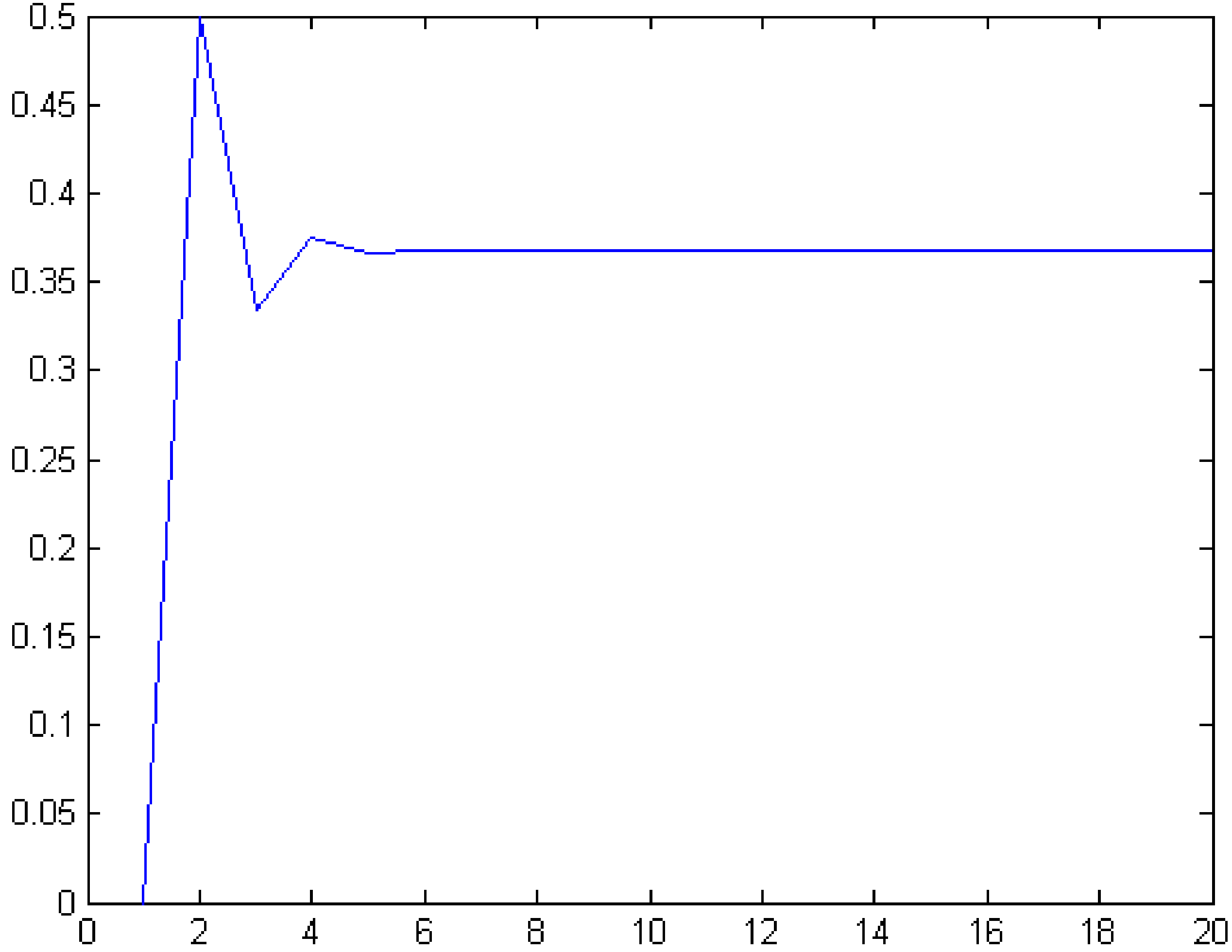
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \sum_{i=0}^{\infty} \frac{x^i}{i!}.$$

Probability of no one getting their backpack back is

$$\sum_{i=2}^{n} (-1)^i \frac{1}{i!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}.$$

which is approximation to e^{-1}, by substituting -1 for x in the power series and stopping at $i = n$.

n	$\sum_{i=0}^{n} (-1)^i \frac{1}{i!}$
1	0.000000000000
2	0.500000000000
3	0.333333333333
4	0.375000000000
5	0.366666666667
6	0.368055555556
7	0.367857142857
8	0.367881944444
9	0.367879188713
10	0.367879464286
11	0.367879439234
12	0.367879441321
13	0.367879441161
14	0.367879441172
15	0.367879441171
16	0.367879441171
17	0.367879441171
18	0.367879441171
19	0.367879441171
20	0.367879441171



Unions and Intersections

- The Probability of a Union of Events
- The Principle of Inclusion and Exclusion for Probability
- The Principle of Inclusion and Exclusion for Counting

The Principle of Inclusion and Exclusion for Counting

How many functions from an n-element set N to an m-element set $M = \{y_1, y_2, \ldots, y_m\}$ map nothing to y_1?

Simply $(m-1)^n$.

Because we have $m-1$ choices
of where to map each of our n elements.

How many functions map nothing to a
given k-element subset K of M?

Using same reasoning as above, number of functions that map nothing to a given set K of k elements will be $(m-k)^n$.

(a) How many onto functions are there from an n-element set N to an m-element set M?

(b) How many functions from an n-element set N to an m-element set M map nothing to at least one element of M?

Since there are exactly m^n functions from an n-element set N to an m-element set M

The answer to (b) is, m^n minus the answer to (a)!

(b) How many functions from an n-element set N to an m-element set M map nothing to at least one element of M?

We need an analog of the principle of inclusion and exclusion for the size of a union of m sets.

Because events are sets, get analog simply by changing probabilities of events E_i to sizes of sets E_i.

Here, set E_i is set of functions that map nothing to element i of set M – that is, event that a function maps nothing to i.)

Principle of inclusion and exclusion for counting:

$$\left|\bigcup_{i=1}^{m} E_i\right| = \sum_{k=1}^{m}(-1)^{k+1} \sum_{\substack{i_1,i_2,\ldots,i_k: \\ 1\le i_1<i_2<\ldots<i_k\le m}} |E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k}|$$

Applying this formula to number of functions from N to M that map nothing to at least one element of K gives

$$\left|\bigcup_{i=1}^{m} E_i\right| = \sum_{k=1}^{m}(-1)^{k+1} \sum_{\substack{i_1,i_2,\ldots,i_k:\\ 1\le i_1<i_2<\ldots<i_k\le m}} |E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_k}|$$

$$= \sum_{k=1}^{m}(-1)^{k+1} \boxed{\binom{m}{k}(m-k)^n}$$

where $|E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_k}|$ is number of functions that map nothing to k-element set $K = \{i_1, i_2, \ldots, i_k\}$.

$\binom{m}{k}$ is number of ways to pick subset K

For fixed K, number of these functions is $(m-k)^n$.

Total number of functions from N to M is m^n.

Thus, number of *onto* functions is

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

$$= \sum_{k=0}^{m} (-1)^{k} \binom{m}{k} (m-k)^n$$

because $\binom{m}{0} = 1$, $(m-0)^n$ is m^n, and $-(-1)^{k+1} = (-1)^k$.

Theorem 5.4:

The number of functions from an n-element set onto an m element set is $\sum_{k=0}^{m} (-1)^{k} \binom{m}{k} (m-k)^n.$