The Inclusion Exclusion Principle and Its More General Version

Stewart Weiss

June 28, 2009

1 Introduction

The Inclusion-Exclusion Principle is typically seen in the context of combinatorics or probability theory. In combinatorics, it is usually stated something like the following:

Theorem 1 (Combinatorial Inclusion-Exclusion Principle)**.** *Let $A_1, A_2, \ldots, A_n$ be finite sets. Then*

$$\begin{aligned}\left|\bigcup_{i=1}^{n} A_i\right| &= \sum_{i_1=1}^{n} |A_{i_1}| - \sum_{i_1=1}^{n-1}\sum_{i_2=i_1+1}^{n} |A_{i_1}\cap A_{i_2}| + \sum_{i_1=1}^{n-2}\sum_{i_2=i_1+1}^{n-1}\sum_{i_3=i_2+1}^{n} |A_{i_1}\cap A_{i_2}\cap A_{i_3}| - \\ &\quad \cdots + \sum_{i_1=1}^{n-k+1}\cdots\sum_{i_k=i_{k-1}+1}^{n} (-1)^{k-1}\,|A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}| + \\ &\quad \cdots + (-1)^{n-1}\,|A_1\cap\cdots\cap A_n|. \end{aligned} \tag{1}$$

To illustrate, when $n = 3$, this reduces to

$$\left|\bigcup_{i=1}^{3} A_i\right| = \sum_{i=1}^{3} |A_i| - \sum_{i=1}^{2}\sum_{j=i+1}^{3} |A_i \cap A_j| + |A_1\cap A_2\cap A_3|.$$

Eq. (1) is more succinctly written as

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{S\subseteq U} (-1)^{|S|-1} \bigcap_{j\in S} A_j \tag{2}$$

where $U = \{1, 2, \ldots, n\}$.

Example. Suppose we ask how many integers less than 120 are relatively prime to it. Since 120 has prime factors 2, 3, and 5, any number that is relatively prime to it is not a multiple of any of these. We can solve this using the *Inclusion-Exclusion Principle* as follows. Let $A = \{1, 2, \ldots, 120\}$, the set of integers from 1 to 120. Let S_1 be the multiples of 2 in A, S_2 be the multiples of 3 in A, and S_3 be the multiples of 5 in A. The set of numbers relatively prime to 120 are the numbers not in $S_1 \cup S_2 \cup S_3$, so there are $120 - |S_1 \cup S_2 \cup S_3|$ many of them. By the theorem,

$$\left|\bigcup_{i=1}^{3} S_i\right| = \sum_{i=1}^{3} |S_i| - \sum_{i=1}^{2}\sum_{j=i+1}^{3} |S_i \cap S_j| + |S_1\cap S_2\cap S_3|$$

A set such as S_1 obviously has $120/2 = 60$ elements. Since a set such as $S_1 \cup S_2$ is the set of numbers that are multiples of both 2 and 3, it is the set of multiples of 6, and has $120/6 = 20$ elements. Applying this idea to the formula, we get

$$
\begin{aligned}
\left|\bigcup_{i=1}^{3} S_i\right| &= \sum_{i=1}^{3}|S_i| - \sum_{i=1}^{2}\sum_{j=i+1}^{3}|S_i \cap S_j| + |S_1 \cap S_2 \cap S_3| \\
&= (60+40+24) - (20+12+8) + 4 \\
&= 124 - 40 + 4 \\
&= 88
\end{aligned}
$$

so there are $120 - 88 = 32$ numbers less than 120 that are relatively prime to it.

2 Generalized Inclusion-Exclusion Principle

The Inclusion-Exclusion Principle actually has a more general form, which can be used to derive the probabilistic and combinatorial versions. This general form, however, is more broadly applicable (which is why it is "more general.") It follows.

Theorem 2. *Let S be a set of N elements and let g and f be functions from the powerset of S into the real numbers with the property that, for any subset $A \subseteq S$*

$$g(A) = \sum_{S' \subseteq A} f(S').$$

Then

$$f(A) = \sum_{S' \subseteq A} (-1)^{|A|-|S'|} g(S').$$

Remark. The theorem takes a while to absorb. Restating it informally, what it says is that if there are two functions f and g, each of which is defined on subsets of a set S, having the property that $g(A)$ is always the sum of the function f applied to every subset of A, including A itself, then it turns out that $f(A)$ can also be expressed as an alternating sum of the function g applied to every subset of A, where what determines whether or not $g(S')$ is added or subtracted from the sum is whether it has an even or an odd number of elements.

Proof. Assume that f and g are functions from the powerset of S into the reals and that, for any subset $A \subseteq S$

$$g(A) = \sum_{S' \subseteq A} f(S'). \tag{3}$$

We show by induction on the cardinality of A that

$$f(A) = \sum_{S' \subseteq A} (-1)^{|A|-|S'|} g(S'). \tag{4}$$

To start, let $|A| = 0$. Then A is the empty set and has no subsets other than itself. Since the premise is true for this A, $g(\emptyset) = f(\emptyset)$. Therefore, $f(\emptyset) = g(\emptyset) = (-1)^{|\emptyset|-|\emptyset|}g(\emptyset)$, so the conclusion is true when $|A| = 0$.

We now assume that the hypothesis is true for all sets B such that $|B| < n$, and we let A be a set of cardinality n. We need to show that if

$$g(A) = \sum_{S' \subseteq A} f(S') \tag{5}$$

then Eq. (4) is true. Let us rewrite (5) as follows:

$$g(A) = \sum_{S' \subset A} f(S') + f(A).$$

Any set S' such that $S' \subset A$ is strictly smaller than A, so $|S'| < n$ and the induction hypothesis applies to it. Therefore for each such S',

$$f(S') = \sum_{T \subseteq S'} (-1)^{|S'|-|T|} g(T) \tag{6}$$

Substituting Eq. (6) into (5), we get

$$g(A) = \sum_{S' \subset A} \sum_{T \subseteq S'} (-1)^{|S'|-|T|} g(T) + f(A). \tag{7}$$

Subtracting and rearranging Eq. (7), we get

$$\begin{aligned} f(A) &= g(A) - \sum_{S' \subset A} \sum_{T \subseteq S'} (-1)^{|S'|-|T|} g(T) \\ &= \sum_{S' \subset A} \sum_{T \subseteq S'} (-1)^{1+|S'|-|T|} g(T) + g(A). \end{aligned} \tag{8}$$

Let T be a fixed subset in the sum and assume that $|T| = r$. Since $T \subset A$ and $|A| = n$, $r < n$. Although T occurs only once in the inner sum $\sum_{T \subseteq S'} (-1)^{1+|S'|-|T|} g(T)$, it may occur multiple times in the entire expression, because it may be a subset of multiple subsets $S' \subset A$. We now do some counting.

There are $n - r$ elements of A that are not in T. Each of these may be independently included or not in a subset containing T. Therefore, there are $2^{(n-r)}$ subsets of A that contain T. One of these is A itself. Since we are enumerating only the proper subsets of A, we have to exclude A from the count, so there are $2^{(n-r)} - 1$ proper subsets S' such that $T \subseteq S'$.

Claim 3. Let A be a set such that $|A| = n$. Then A has $2^{n/2}$ subsets of even cardinality and $2^{n/2}$ subsets of odd cardinality.

Proof. This is easy to prove by induction on the cardinality of A. □

Claim 4. Let A be a set such that $|A| = n$ and let T be a subset of A such that $|T| = r$. Then $A - T$ has $2^{(n-r)/2}$ subsets of even cardinality and $2^{(n-r)/2}$ subsets of odd cardinality.

Proof. This follows from Claim 3. □

Claim 5. Let A be a set such that $|A| = n$ and let T be a subset of A such that $|T| = r$. Then there are $2^{(n-r)/2}$ subsets of A of even cardinality that contain T, and $2^{(n-r)/2}$ subsets of A of odd cardinality that contain T.

Proof. There is a one-to-one correspondence between the subsets of $A - T$ and the subsets of A containing T : for each subset S' of $A - T$, there is a unique $S'_T = S' \cup T$, and for each S'_T containing T, the set $S'_T - T$ is a unique subset of $A - T$. By Claim 4, the claim is proved. □

Claim 6. Let A be a set such that $|A| = n$ and let T be a subset of A such that $|T| = r$. Let $M = 2^{(n-r)}$. If n is odd, then there are $M/2$ *proper subsets* of A of even cardinality and $(M/2) - 1$ *proper subsets* of A of odd cardinality containing T. If n is even, then there are $M/2$ *proper subsets* of A of odd cardinality and $(M/2) - 1$ *proper subsets* of A of even cardinality containing T.

Proof. The proper subsets of A do not include A. Since the number of even and odd subsets is equal when A is included, then if n is odd, there is one fewer odd subset than even ones, and if n is even, there is one fewer even subset than odd. □

Consider the coefficient $(-1)^{1+|S'|-|T|}$ of $g(T)$ in Eq. (8). It is either $+1$ or -1. For a fixed subset T, this depends only on whether $|S'|$ is even or odd. Claim 6 established that, for fixed T, the number of subsets S' containing T and of even cardinality differs from the number of those with odd cardinality by exactly one. Therefore, all but one of the terms

$$(-1)^{1+|S'|-|T|} g(T)$$

will cancel out, leaving either $+g(T)$ or $-g(T)$. Suppose r is even, then

$$(-1)^{1+|S'|-|T|}g(T) = (-1)^{1+|S'|}g(T).$$

If $|S'|$ is even, $(-1)^{1+|S'|} = -1$ and if $|S'|$ is odd, then $(-1)^{1+|S'|} = 1$. Therefore, if the number of even subsets S' is one greater than the number of odd subsets S', then the sum of the coefficients of all terms containing $g(T)$ is -1, and if the number of even subsets is one less, then the coefficient is $+1$. Claim 6 shows that when n is odd, there is one more even subset than odd subset and so the sum of the coefficients is -1. When n is even, there is one more odd subset than even, and so the sum of the coefficients is 1. Therefore, when r is even and n is odd, the coefficient of $g(T)$ is -1 and when r is even and$n$ is even, the coefficient of $g(T)$ is $+1$.

On the other hand, when r is odd,

$$(-1)^{1+|S'|-|T|}g(T) = (-1)^{|S'|}g(T)$$

and the signs reverse, so when n is even and r is odd, the coefficient of $g(T)$ is -1 and when n is odd and r is odd, the coefficient of $g(T)$ is 1. In short, the coefficient of $g(T)$ is $(-1)^{|A|-|T|}$. Therefore

$$\begin{aligned} f(A) &= \sum_{S' \subset A} \sum_{T \subseteq S'} (-1)^{1+|S'|-|T|}g(T) + g(A) \\ &= \sum_{T \subset A} (-1)^{|A|-|T|}g(T) + g(A) \\ &= \sum_{T \subseteq A} (-1)^{|A|-|T|}g(T). \end{aligned}$$

This is exactly what we needed to prove, so by the axiom of induction, the theorem is proved. $\square$

3 The Combinatorial Version as an Instance of the General Version

We now show that Theorem 1 is an instance of Theorem 2. We introduce some notation.

Notation. If A and B are sets, then $A - B$ is the set of all elements in A that are *not* in B. This is sometimes called the *set difference* between A and B, or the *complement* of B in A. For example, $\{1, 2, 3, 4, 5\} - \{3, 4, 6, 7\} = \{1, 2, 5\}$. When all of the sets being considered are subsets of some larger, fixed universe of discourse U, then we define, for each set S, the *complement* of S in U, denoted $\overline{S}$, by $\overline{S} = U - S$. When the universe is understood, for simplicity we call $\overline{S}$ the *complement* of S.

Suppose that $A_1, A_2, A_3, \ldots, A_n$ is a collection of finite sets. Let $A = \bigcup A_i$. Let $U = \{1, 2, 3, \ldots, n\}$. The set U will act as our universe. Let S be a subset of U and define the function

$$F(S) = \begin{cases} \bigcap_{j \in \overline{S}} A_j - \bigcup_{j \in S} A_j & S \neq U \\ \emptyset & S = U \end{cases}$$

For example, let $n = 6$ and let $S = \{3, 4, 5\}$. Then $\overline{S} = \{1, 2, 6\}$ and $F(S) = (A_1 \cap A_2 \cap A_6) - (A_3 \cup A_4 \cup A_5)$. In other words, $F(S)$ is the set of all elements, each of which is in A_1, A_2, and A_6 but not in any of A_3, A_4, or A_5.

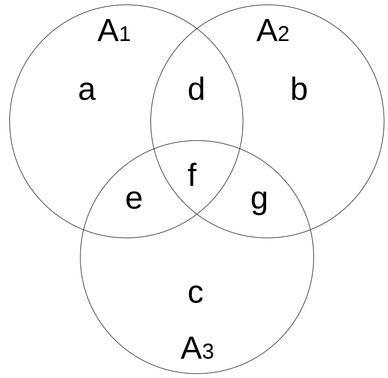

Figure 1: Example

If it seems at first that $F()$ is an arcane function with little practical value, then consider the following example, illustrated by Figure 1. Let $A_1 = \{a, d, e, f\}$, let $A_2 = \{b, d, f, g\}$, and let $A_3 = \{c, e, f, g\}$. Let $U = \{1, 2, 3\}$. Consider the subset $S = \{1, 3\}$ of U. For this subset, $F(S) = A_2 - (A_1 \cup A_3) = \{b, d, f, g\} - \{a, d, e, f, c, g\} = \{b\}$. Notice that $\{b\}$ is exactly that part of A_2 not in any other set. Note that $F(\{1\}) = (A_2 \cap A_3) - A_1 = \{f, g\} - \{a, d, e, f\} = \{g\}$, which is the element in another distinct region of the figure. The function $F()$ is actually creating a partition of the set of all elements, with each cell of that partition corresponding to each one of the 2^n subsets of U. This is the essence of the next two lemmas.

Lemma 7. *If $S_1 \subseteq U$ and $S_2 \subseteq U$ and $S_1 \neq S_2$, then $F(S_1) \cap F(S_2) = \emptyset$.*

Proof. Since $S_1 \neq S_2$, there is a number m such that $1 \leq m \leq n$ and either $m \in S_1$ and $m \notin S_2$, or $m \in S_2$ and $m \notin S_1$. Without loss of generality, we can assume that $m \in S_1$ and $m \notin S_2$. Suppose that the theorem is false and that $F(S_1) \cap F(S_2) \neq \emptyset$ Let $x \in F(S_1) \cap F(S_2)$. Since $x \in F(S_1)$,

$$x \in \bigcap_{j \in \overline{S_1}} A_j - \bigcup_{j \in S_1} A_j.$$

This implies that

$$x \notin \bigcup_{j \in S_1} A_j$$

Since $m \in S_1$, this implies in turn that

$$x \notin A_m. \tag{9}$$

Since $x \in F(S_2)$,

$$x \in \bigcap_{j \in \overline{S_2}} A_j - \bigcup_{j \in S_2} A_j$$

which implies, in particular, that

$$x \in \bigcap_{j \in \overline{S_2}} A_j.$$

Since$m \notin S_2$, $m \in \overline{S_2}$ and so $x \in A_m$. But this contradicts (9) above, so we have reached a contradiction. Therefore $F(S_1) \cap F(S_2) = \emptyset$ and the lemma is proved. □

Lemma 8. *For each $x \in A$, there exists a set $S \subseteq U$ such that $x \in F(S)$.*

Proof. Let $x \in A$. Then x is in at least one of the sets A_i. Let S be the set of all numbers j such that $x \notin A_j$. If x is a member of all sets A_j then $S = \emptyset$. Since $x \notin A_j$ for all $j \in S$,

$$x \notin \bigcup_{j \in S} A_j \tag{10}$$

Suppose k is a number such that $k \notin S$. Consider the set A_k. Since $k \notin S$, A_k must contain x. This is true for all of the sets A_k for which $k \notin S$. In other words, for all $k \in \overline{S}$, $x \in A_k$. Therefore

$$x \in \bigcap_{k \in \overline{S}} A_k \tag{11}$$

Eqs. (10) and (11) combined show that

$$x \in \bigcap_{j \in \overline{S}} A_j - \bigcup_{j \in S} A_j.$$

which implies that

$$x \in F(S)$$

proving the lemma. $\square$

Lemmas 7 and 8 together show that, for each element x of A, there is a unique $S \subseteq U$ such that $x \in F(S)$. Therefore, the sum of the cardinalities of the sets $F(S)$ ranging over all subsets $S \subseteq U$ equals the number of elements in A, i.e.,

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{S \subseteq U} |F(S)|.$$

Let f be defined by

$$f(S) = \Big| F(S) \Big|. \tag{12}$$

and let g be defined by

$$g(S) = \begin{cases} \left| \bigcap_{j \in \overline{S}} A_j \right| & S \neq U \\ \left| \bigcup_{j=1}^{n} A_j \right| & S = U \end{cases} \tag{13}$$

Then

$$g(U) = \sum_{S \subseteq U} f(S).$$

By Theorem 2,

$$f(U) = \sum_{S \subseteq U} (-1)^{|U|-|S|} g(S)$$

Since $f(U) = |F(U)| = 0$,

$$\begin{aligned} 0 &= \sum_{S \subseteq U} (-1)^{|U|-|S|} g(S) \\ &= \sum_{S \subsetneq U} (-1)^{|U|-|S|} g(S) + g(U) \end{aligned}$$

which implies that

$$g(U) = \sum_{S \subsetneq U} (-1)^{|U|-|S|-1} g(S)$$

Substituting (),

$$\left|\bigcup_{j=1}^{n} A_j\right| = \sum_{S \subsetneq U} (-1)^{|U|-|S|-1} g(S) \tag{14}$$

The right-hand side of (14) is a sum over all subsets of U except for U itself. For any subset $S \neq U$, let $T = \overline{S}$. Then $S = \overline{T}$. Since $|\overline{T}| = |U| - |T|$, $|U| - |\overline{T}| - 1 = |T| - 1$. Also, $g(\overline{T}) = \left|\bigcap_{j \in T} A_j\right|$. Therefore

$$\begin{aligned}
(-1)^{|U|-|S|-1} g(S) &= (-1)^{|U|-|\overline{T}|-1} g(\overline{T}) \\
&= (-1)^{|T|-1} \left|\bigcap_{j \in T} A_j\right|
\end{aligned}$$

If Eq. (14) is true when the right-hand side ranges over all subsets of U except U, then if we replace the subsets by their complements, Eq. (14) will be true when the right-hand side ranges over all subsets except $\overline{U} = \emptyset$. Therefore,

$$\left|\bigcup_{j=1}^{n} A_j\right| = \sum_{\emptyset \neq T \subseteq U} (-1)^{|T|-1} \left|\bigcap_{j \in T} A_j\right| \tag{15}$$

and replacing the variable T by S we see that this is identical to Eq. (2). This proves that the combinatorial version of the IEP is just a special case of Theorem 2.