Inclusion-Exclusion principle

Jonathan L.F. King
University of Florida, Gainesville FL 32611-2082, USA
squash@ufl.edu
Webpage http://squash.1gainesville.com/

12 May, 2024 (at *13:16*)

Prolegomenon. Consider a finite set, Ω, of ***tokens***. and a finite indexing set $\mathcal{N}$; let $N := {}^{\#}\mathcal{N}$. For each $j \in \mathcal{N}$ we have a ***patch***, a subset $A_j \subset \Omega$, and its complement $V_j := \Omega \smallsetminus A_j$. Our goal is an expression for the cardinality of the union

$$\mathbf{U} := \bigcup_{j\in\mathcal{N}} A_j \,.$$

[N.B: This note uses $^{\#}$ to indicate the cardinality of an index-set, and $|\cdot|$ for the cardinality of a token set.]

Notation. For each ***index-set*** $I \subset \mathcal{N}$, define the ***patch-intersection***

1.1: $$\mathbf{A}_I := \bigcap_{j\in I} A_j \,,$$

[and $\mathbf{V}_I := \bigcap_{j\in I} V_j$] and note[♡1] $\mathbf{A}_\varnothing$ is all of Ω.

Finally, let $\mathcal{C}_k$ comprise those index-sets $I \subset \mathcal{N}$ with ${}^{\#}I = k$; this, for $k = 0, 1, \dots, N$. Easily

1.2: $$|\mathcal{C}_k| = \binom{N}{k} \,.$$

2: Inclusion-Exclusion Lemma. *With notation from above:*

2a: $$\big|\Omega \smallsetminus \mathbf{U}\big| = \sum_{k=0}^{N} \Big[[-1]^k \cdot \sum_{I\,:\,I\in\mathcal{C}_k} |\mathbf{A}_I|\Big] \,.$$

Alternatively,

2b: $$\big|\mathbf{U}\big| = \sum_{k=1}^{N} \Big[[-1]^{k-1} \cdot \sum_{I\,:\,I\in\mathcal{C}_k} |\mathbf{A}_I|\Big] \,. \quad \diamondsuit$$

Pf (2a) ⇒ (2b). Adding the RhSes causes cancellation, with only the k=0 term remaining. So

$$|\Omega \smallsetminus \mathbf{U}| + \mathrm{RhS}(2b) = [-1]^0 \cdot |\mathbf{A}_\varnothing| \overset{\text{note}}{=\!=} |\Omega| \,.$$

So RhS(2b) equals $|\mathbf{U}|$. ♦

[♡1] For index-sets $I, J \subset \mathcal{N}$, by definition $\mathbf{A}_I \cap \mathbf{A}_J = \mathbf{A}_{I\cup J}$. In particular, $\mathbf{A}_I \cap \mathbf{A}_\varnothing$ equals $\mathbf{A}_I$. I.e, $\mathbf{A}_\varnothing$ is the identity element for intersection, on the powerset of Ω. So $\mathbf{A}_\varnothing$ is Ω.

To establish (2a), let's prove something a little stronger. Each subset $S \subset \Omega$ yields a function $\mathbf{1}_S : \Omega \to \{0,1\}$, the "***indicator function*** of S",

$$\mathbf{1}_S(x) := \begin{Bmatrix} 1 & \text{when } x \in S \\ 0 & \text{when } x \in \Omega \smallsetminus S \end{Bmatrix} .$$

Indicator fncs allow us to restate intersection ITOf multiplication: *For a set I of indicies,*

1.1′: $$\mathbf{1}_{\mathbf{A}_I} = \prod_{j\in I} \mathbf{1}_{A_j} \,.$$

Let's strengthen (2a) to equality of *functions*,

3.1: $$\mathbf{1}_{\Omega\smallsetminus\mathbf{U}} = \sum_{k=0}^{N} \Big[[-1]^k \cdot \sum_{I\,:\,I\in\mathcal{C}_k} \mathbf{1}_{\mathbf{A}_I}\Big] \,.$$

Proof of (3.1). The RhS(3.1) equals

$*$: $$\prod_{j\in\mathcal{N}} [1 - \mathbf{1}_{A_j}] \overset{\text{note}}{=\!=} \prod_{j\in\mathcal{N}} \mathbf{1}_{V_j} \,.$$

After all, the LhS($*$) product expands to the sum, over all subsets $I \subset \mathcal{N}$, of $\prod_{j\in I} [-\mathbf{1}_{A_j}]$. This latter product, letting $k := {}^{\#}I$, equals $[-1]^k \cdot \mathbf{1}_{\mathbf{A}_I}$, courtesy (1.1′).

By De Morgan's law, $\Omega \smallsetminus \mathbf{U} = \bigcap_{j\in\mathcal{N}} V_j$. So LhS(3.1) is $\mathbf{1}_{\mathbf{V}_\mathcal{N}}$. Hence (3.1) boils down to the triviality that

$$\mathbf{1}_{\mathbf{V}_\mathcal{N}} = \prod_{j\in\mathcal{N}} \mathbf{1}_{V_j} \,. \quad ♦$$

4: *Rem.* We can rewrite (2a) in simpler [but often less convenient] form, ITOf summing over all subsets of $\mathcal{N}$:

4a: $$\big|\Omega \smallsetminus \mathbf{U}\big| = \sum_{I\,:\,I\subset\mathcal{N}} \big[[-1]^{\#I} \cdot |\mathbf{A}_I|\big] \,. \quad \square$$

Incl-Excl Examples

5: Counting limited candy. *The store sells jelly-$\mathcal{B}$eans and $\mathcal{C}$hocolate squares and $\mathcal{D}$ates. Mom allows you a total of 20 candies.*

Alas!, the store only has $8\mathcal{B}$ and $5\mathcal{C}$ and $13\mathcal{D}$. Stars-and-Bars counts how to pick out of multiset $\{\infty\mathcal{B}, \infty\mathcal{C}, \infty\mathcal{D}\}$. The relevant multiset is $\{8\mathcal{B}, 5\mathcal{C}, 13\mathcal{D}\}$; so how do we count? ♢

Candy soln. Let Ω be the set of natnum triples $(\mathcal{B},\mathcal{C},\mathcal{D})$ with $\mathcal{B}+\mathcal{C}+\mathcal{D}=20$. We'll count the "***good***" $[\mathcal{B}\le 8\ \&\ \mathcal{C}\le 5\ \&\ \mathcal{D}\le 13]$ triples, using Incl-Excl.

Let $A_{\mathcal{B}}$ be the set of natnum-triples that are "***Awful***" because $\mathcal{B}>8$. Hence,

$$|A_{\mathcal{B}}| \stackrel{Why?}{=\!=} \left[\!\!\left[\begin{matrix}3\\ 20-[8+1]\end{matrix}\right]\!\!\right] = \binom{2+11}{2} = 78.$$

So $|A_{\mathcal{C}}| = \left[\!\left[\begin{smallmatrix}3\\ 20-[5+1]\end{smallmatrix}\right]\!\right] = \binom{2+14}{2} = 120$, and $|A_{\mathcal{D}}| = 28$.

For *pairwise* intersections

$$|A_{\mathcal{B}}\cap A_{\mathcal{C}}| \stackrel{Why?}{=\!=} \left[\!\!\left[\begin{matrix}3\\ 20-[8+5+2]\end{matrix}\right]\!\!\right] = \binom{2+5}{2} = 21.$$

Also, $|A_{\mathcal{B}}\cap A_{\mathcal{D}}| = \left[\!\left[\begin{smallmatrix}3\\ 20-[8+13+2]\end{smallmatrix}\right]\!\right] = \left[\!\left[\begin{smallmatrix}3\\ \text{negative}\end{smallmatrix}\right]\!\right] \stackrel{Why?}{=\!=} 0$, and $|A_{\mathcal{C}}\cap A_{\mathcal{D}}| = \left[\!\left[\begin{smallmatrix}3\\ 20-[5+13+2]\end{smallmatrix}\right]\!\right] = \left[\!\left[\begin{smallmatrix}3\\ 0\end{smallmatrix}\right]\!\right] = 1$.

For the sole *three-fold* intersection

$$|A_{\mathcal{B}}\cap A_{\mathcal{C}}\cap A_{\mathcal{D}}| = \left[\!\!\left[\begin{matrix}3\\ 20-[8+5+13+3]\end{matrix}\right]\!\!\right] = \left[\!\!\left[\begin{matrix}3\\ \text{neg}\end{matrix}\right]\!\!\right] = 0.$$

Since $\left[\!\left[\begin{smallmatrix}3\\ 20\end{smallmatrix}\right]\!\right] = 231$, the number of good triples is

$$\begin{aligned} &|\Omega| - \Big(|A_{\mathcal{B}}| + |A_{\mathcal{C}}| + |A_{\mathcal{D}}|\Big) \\ &\quad + \Big(|A_{\mathcal{B}}\cap A_{\mathcal{C}}| + |A_{\mathcal{B}}\cap A_{\mathcal{D}}| + |A_{\mathcal{C}}\cap A_{\mathcal{D}}|\Big) \\ &\quad - |A_{\mathcal{B}}\cap A_{\mathcal{C}}\cap A_{\mathcal{D}}| \\ &= 231 - [78+120+28] + [21+0+1] - 0. \end{aligned}$$

This equals 27. ♦

Doubting Thomas. Here are the 27 good triples:

```
(2 5 13) (3 4 13) (3 5 12) (4 3 13) (4 4 12) (4 5 11)
(5 2 13) (5 3 12) (5 4 11) (5 5 10) (6 1 13) (6 2 12)
(6 3 11) (6 4 10) (6 5  9) (7 0 13) (7 1 12) (7 2 11)
(7 3 10) (7 4  9) (7 5  8) (8 0 12) (8 1 11) (8 2 10)
(8 3  9) (8 4  8) (8 5  7)
```

□

6: *Cardinality independence*. In some combinatorial applications, cardinality $|\mathbf{A}_I|$ depends only on the number of patches being intersected. In that instance, let $F(k)$ be $|\mathbf{A}_I|$, for each and every index-set I satisfying $^{\#}I = k$. So rewrite (2a) as

6a:
$$\begin{aligned} |\Omega \smallsetminus \mathbf{U}| &= \sum_{k=0}^{N} \Big[[-1]^k \cdot \tbinom{N}{k} \cdot F(k)\Big] \\ &= \sum_{k=0}^{\infty} \Big[[-1]^k \cdot \tbinom{N}{k} \cdot F(k)\Big], \end{aligned}$$

where, for $k>N$, our $\binom{N}{k} := \frac{[\![N\downarrow k]\!]}{k!}$ is zero. □

7: Probability of getting your own hat. *The N guests leaving your party grab their hats at random from your dark closet. What does $\Pr(N)$, the probability that no one gets his own hat, tend to as $N \nearrow \infty$?* ◊

Hat soln. For a subset I of your guests, the probability that each person in I took his own hat is $\mathcal{P}_I := [N-k]!/N!$, where $k := {}^{\#}I$. As this probability only depends on k, we will also call it $\mathcal{P}_k$. By the principle of inclusion/exclusion our $\Pr(N)$ equals

$$\mathcal{P}_\varnothing - \sum_{I:\,{}^{\#}I=1}\mathcal{P}_I + \sum_{I:\,{}^{\#}I=2}\mathcal{P}_I - \sum_{I:\,{}^{\#}I=3}\mathcal{P}_I + \ldots + [-1]^N \sum_{I:\,{}^{\#}I=N}\mathcal{P}_I$$

$$= 1 - \binom{N}{1}\mathcal{P}_1 + \binom{N}{2}\mathcal{P}_2 - \binom{N}{3}\mathcal{P}_3 + \ldots + [-1]^N\binom{N}{N}\mathcal{P}_N$$

$$= \sum_{k=0}^{N}[-1]^k \cdot \binom{N}{k} \cdot \frac{[N-k]!}{N!} = \sum_{k=0}^{N}\frac{[-1]^k}{k!}.$$

This last is the first $N+1$ terms of the Taylor series for e^{-1}. Thus $\lim_{n\to\infty}\Pr(n)$ equals $1/\mathrm{e} \approx 0.368$. ♦

Derangements. A ***derangement*** of an N-set is a *fixed-point free* permutation of that set; let $\mathbb{D}_N$ be the *set* of derangements, and $\boldsymbol{\delta}_N := |\mathbb{D}_N|$ the number of derangements. Thus

7a: $$\boldsymbol{\delta}_N = N! \cdot \sum_{k=0}^{N}\frac{[-1]^k}{k!}.$$

is a restatement of the above probability.

Counting perms with fixed-pts. For $N = 0, 1, 2, \ldots$, we define $\mathrm{PFix}_N(\mathbf{f})$ as the Number of N-Permutations with *precisely* $\mathbf{f}$ many Fixed-points. So $\mathbf{f}<0$ or $\mathbf{f}>N$ makes $\mathrm{PFix}_N(\mathbf{f}) = 0$. [By defn, $\mathrm{PFix}_N(0) = \boldsymbol{\delta}_N$.] □

7b: Corollary. *Each $\mathbf{f}$ in $[0\,..\,N]$ satisfies*

†f: $$\mathrm{PFix}_N(\mathbf{f}) = \frac{N!}{\mathbf{f}!} \cdot \sum_{k=0}^{N-\mathbf{f}}\frac{[-1]^k}{k!}$$

Thus, the asymptotic probability of $\mathbf{f}$ fixed-pts is

‡: $$\lim_{N\to\infty}\frac{\mathrm{PFix}_N(\mathbf{f})}{N!} = \frac{1/\mathrm{e}}{\mathbf{f}!}.$$ ◊

Proof. WLOG, $\mathbf{f} = 3$. The number of permutations of $[1\,..\,N]$ that fix, say, points $2, 6, 9$ is the number of derangements of $[1\,..\,N] \smallsetminus \{2,6,9\}$; so there are $\boldsymbol{\delta}_{N-3}$ such perms.

Token-set $[1\,..\,N]$ has $\binom{N}{3}$ subsets of size 3. Thus

$$\mathrm{PFix}_N(3) = \binom{N}{3} \cdot \boldsymbol{\delta}_{N-3} \overset{\text{simplify}}{=\!=\!=} \mathrm{RhS}(\dagger_3).$$ ♦

Prelim. Below, sets $\mathfrak{D}$ (Domain) and $\mathfrak{C}$ (Codomain) have cardinalities $\mathsf{D} := |\mathfrak{D}|$ and $\mathsf{C} := |\mathfrak{C}|$; both finite. Thus $\mathfrak{C}^{\mathfrak{D}}$, the set of fncs $\mathfrak{D}\to\mathfrak{C}$, has cardinality C^{D}. Easily:

$$*: \qquad [\text{The \# of injections } \mathfrak{D}\to\mathfrak{C}] = [\![\mathsf{C}\downarrow\mathsf{D}]\!].$$

Let's compute $\mathrm{Sur}(\mathsf{D},\mathsf{C})$, the number of *surjections*. □

8a: Counting surjective fncs. *With notation from above*

$$\dagger: \qquad \mathrm{Sur}(\mathsf{D},\mathsf{C}) = \sum_{k=0}^{\mathsf{C}} [-1]^k \cdot \binom{\mathsf{C}}{k} \cdot [\mathsf{C} - k]^{\mathsf{D}}. \qquad \lozenge$$

Sur. For point $y \in \mathfrak{C}$, let A_y comprise those functions $h()$ which *Avoid* y; i.e, $\mathrm{Range}(h) \not\ni y$. Thus

$$\ddagger: \qquad \mathfrak{C}^{\mathfrak{D}} \smallsetminus \Big[\bigcup_{y\in\mathsf{C}} A_y\Big]$$

is the *set* of surjections.

For $I \subset \mathfrak{C}$, let $\mathbf{A}_I$ comprise those fncs which miss *each* member of I. With $k := {}^{\#}I$, then,

$$\mathbf{A}_I = \{h \in \mathfrak{C}^{\mathfrak{D}} \mid \mathrm{Range}(h) \sqcap I\} \quad \textit{and} \quad |\mathbf{A}_I| = [\mathsf{C} - k]^{\mathsf{D}}.$$

The number of subsets $I \subset \mathfrak{C}$ with ${}^{\#}I = k$ is $\binom{\mathsf{C}}{k}$. Consequently, Inclusion-Exclusion yields (†). ♦

When $\mathsf{D}<\mathsf{C}$. There are *no* surjections, when $\mathsf{D}<\mathsf{C}$. As a (†)-example, $\mathrm{Sur}(2,3)$ equals

$$\binom{3}{0}\cdot 3^2 - \binom{3}{1}\cdot 2^2 + \binom{3}{2}\cdot 1^2 - \binom{3}{3}\cdot 0^2$$
$$= 1\cdot 9 - 3\cdot 4 + 3\cdot 1 - 1\cdot 0 = 9 - 12 + 3,$$

which indeed equals zero. □

[*A Curious Corollary* of Counting sur-fncs.]

8b: A Curious Corollary. *For $N = 0, 1, 2, \ldots$*

$$\pounds_N: \qquad N! = \sum_{k=0}^{N} [-1]^k \cdot \binom{N}{k} \cdot [N - k]^N. \qquad \lozenge$$

Proof. When $|\mathfrak{D}| = |\mathfrak{C}| =: N$, then we can identify $\mathfrak{D}$ with $\mathfrak{C}$ and view each surjection as a permutation. There are $N!$ permutations. And $\mathrm{RhS}(\pounds_N)$ equals $\mathrm{RhS}(\dagger)$ when $\mathsf{D} = \mathsf{C} = \mathbb{N}$. ♦

When $|\mathfrak{D}| = |\mathfrak{C}| = 3$. Computing, $\mathrm{Sur}(3,3)$ equals

$$\binom{3}{0}\cdot 3^3 - \binom{3}{1}\cdot 2^3 + \binom{3}{2}\cdot 1^3 - \binom{3}{3}\cdot 0^3$$
$$= 1\cdot 27 - 3\cdot 8 + 3\cdot 1 - 1\cdot 0 = 27 - 24 + 3 = 6,$$

which, happily, equals 3-factorial. □

TwoStirling numbers. For natnums D,C, the number of partitions of a D-set into C many non-void-atoms, is a "***Stirling # of the 2nd kind***", (or ***Stirling partition number***). Here, I'll write it as $\mathcal{S}(\mathsf{D},\mathsf{C})$.

Were the C many atoms *labeled*, then we could view a partition as a surjective [each atom is non-empty] *function* from the D-set into the label-set. Consequently,

$$8c: \qquad \mathcal{S}(\mathsf{D},\mathsf{C}) = \frac{\mathrm{Sur}(\mathsf{D},\mathsf{C})}{\mathsf{C}!} = \sum_{k=0}^{\mathsf{C}} [-1]^k \cdot \frac{[\mathsf{C} - k]^{\mathsf{D}}}{k! \cdot [\mathsf{C} - k]!}$$
$$\underline{\underline{(k,n) \in \mathbb{N}\times\mathbb{N}}} \quad \sum_{k+n=\mathsf{C}} [-1]^k \cdot \frac{n^{\mathsf{D}}}{k! \cdot n!}$$

is the nifty formula we obtain. □

9: Comps avoiding a part. *What is $f(N)$, the number of 3-compositions of N with no PartSize=2?* ♢

Remark. Bona's text asks for $f(12)$. □

Proof. Not yet typed. ♦

Recursive Incl-Excl. This formula appears in *Reduced Recursive Inclusion-exclusion Principle for the probability of union events* by S. G. Chen.

When VPN-ed in to UF, the article is online.

DEFN: For patches $P_1, \ldots, P_K$, let $C(P_1, \ldots, P_K)$ be the cardinality of $\bigcup_{j=1}^{K} P_j$.

Given patches $A_1, \ldots, A_N$, let $\mathbf{U}_k := \bigcup_{j=1}^{k} A_j$; so $\mathbf{U}_0 = \varnothing$. Easily

$$\bigcup_{j=1}^{N} A_j \quad = \quad \bigsqcup_{j=1}^{N} [A_j \smallsetminus \mathbf{U}_{j-1}] \,,$$

writing a union as a *disjoint* union.

Taking $j = 4$ as an example, note

$$\Big|A_4 \smallsetminus \mathbf{U}_3\Big| \;=\; |A_4| - C(A_4 \cap A_1,\, A_4 \cap A_2,\, A_4 \cap A_3)\,.$$

Thus

$$*: \quad C\Big(A_1, \ldots, A_N\Big) = \sum_{j=1}^{N} \Big[|A_j| - C(A_j \cap A_1,\, \ldots,\, A_j \cap A_{j-1})\Big]$$

gives a recursive characterization of $C(\cdot)$, noting that $C(\text{Empty list})$ is *zero*, since it equals the empty-sum, from above.