Symmetric Inclusion-Exclusion

Ira M. Gessel
Brandeis University
Institut Mittag-Leffler

Séminaire Lotharingien de Combinatoire, April 4, 2005
Maison Saint-Bernard, Lucelle, France
Inclusion-exclusion

Suppose that \( A \) and \( B \) are two functions defined on finite sets. Then the following are equivalent:

\[
A(S) = \sum_{T \subseteq S} B(T).
\]

\[
B(S) = \sum_{T \subseteq S} (-1)^{|S| - |T|} A(T).
\]

The special case in which \( A(S) \) and \( B(S) \) depend only on \(|S|\) is especially important: given two sequences \((a_n)\) and \((b_n)\), the following are equivalent:

\[
a_n = \sum_{k=0}^{n} \binom{n}{k} b_k
\]

\[
b_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} a_k
\]
Inclusion-exclusion

Suppose that $A$ and $B$ are two functions defined on finite sets. Then the following are equivalent:

$$A(S) = \sum_{T \subseteq S} B(T).$$

$$B(S) = \sum_{T \subseteq S} (-1)^{|S| - |T|} A(T).$$

The special case in which $A(S)$ and $B(S)$ depend only on $|S|$ is especially important: given two sequences $(a_n)$ and $(b_n)$, the following are equivalent:

$$a_n = \sum_{k=0}^{n} \binom{n}{k} b_k$$

$$b_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} a_k$$
It’s very easy to find pairs of functions (or sequences) with combinatorial interpretations that satisfy these formulas. Just pick $B$ to be anything, and then $A$ will have a combinatorial interpretation.
In terms of exponential generating functions, we have the familiar formulas

\[ a(x) = e^x b(x) \]
\[ b(x) = e^{-x} a(x), \]

where

\[ a(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!} \quad \text{and} \quad b(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!} \]
Now in the first equation let’s replace $B(T)$ with $(-1)^{|T|}B(T)$. We get the symmetric equations

\[
A(S) = \sum_{T \subseteq S} (-1)^{|T|} B(T).
\]

\[
B(S) = \sum_{T \subseteq S} (-1)^{|T|} A(T).
\]

and

\[
a_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} b_k
\]

\[
b_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} a_k
\]
We will call pairs of functions $A$ and $B$ satisfying

$$A(S) = \sum_{T \subseteq S} (-1)^{|T|} B(T).$$

$$B(S) = \sum_{T \subseteq S} (-1)^{|T|} A(T).$$

symmetric inclusion-exclusion function pairs, and we call sequences $(a_n)$ and $(b_n)$ satisfying

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

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

symmetric inclusion-exclusion sequence pairs.
The exponential generating functions for symmetric inclusion-exclusion sequence pairs satisfy the unfamiliar-looking

\[ a(x) = e^x b(-x) \]
\[ b(x) = e^x a(-x) \]
Can we find any symmetric inclusion-exclusion pairs with combinatorial interpretations?
There is a very easy way to convert asymmetric inclusion-exclusion sequence pairs to symmetric inclusion-exclusion sequence pairs. We will illustrate with the special case of the derangement numbers. We take $a_n = n!$, so $b_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} k! = D_n$. We can compute $b_n$ from $a_n$ by using a difference table, in which each number in a row below the first is the number above it to the right minus the number above it to the left.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>6</th>
<th>24</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>18</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>14</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>11</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If we switch the rows and diagonals, we get a “sum table”:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>9</th>
<th>44</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>14</td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where each number in a row greater than the first is the sum of the two numbers above it.
Now let's take our first table and reverse it left-to-right.

```
  120  24  6  2  1  1
  96   18  4  1  0
  78   14  3  1  1
  64   11  2  4  4
  53    9
  44
```
Now each number below the first row is the difference between the number above it to the left minus the number above it to the right. If we switch the rows and diagonals, we get an array with the same property:
Do the $a_i$ and $b_i$ have combinatorial interpretations? Yes: $b_i$ is the number of permutations of $[5]$, in which 1, 2, . . . , $i$ (or any $i$ numbers from $[5]$) are all fixed points and $a_i$ is the number of permutations of $[5]$ in which 1, 2, . . . , $i$ (or any $i$ numbers from $[5]$) are all nonfixed points.
So we can get a symmetric inclusion-exclusion sequence pair by taking a sequence from an asymmetric inclusion-exclusion sequence pair and reversing an initial part of it. But this isn’t satisfactory because

- We can only get *finite* sequences by this method.
- It isn’t very interesting.
So we can get a symmetric inclusion-exclusion sequence pair by taking a sequence from an asymmetric inclusion-exclusion sequence pair and reversing an initial part of it. But this isn’t satisfactory because

- We can only get finite sequences by this method.
- It isn’t very interesting.
Symmetric inclusion-exclusion functions

To get interesting infinite symmetric inclusion-exclusion sequences we must let $a_i$ and $b_i$ be probabilities rather than integers.

We have a probability space containing sets $U_i, i \in \Delta$. (In our examples, $\Delta$ will be $\mathbb{P} = \{1, 2, 3, \ldots \}$.) This means that we have a function $P$ defined on all sets generated from the $U_i$ by complements and finite unions and intersections with the property that if $S$ and $T$ are disjoint then

$$P(S \cup T) = P(S) + P(T).$$

In order for $P$ to be a probability, we also need $P(S) \geq 0$ and $P(\text{universal set}) = 1$, but we don’t really need these properties.
For every subset $S \subseteq \Delta$ we define

$$A(S) = P\left( \bigcap_{i \in S} U_i \right)$$

$$B(S) = P\left( \bigcap_{i \in S} \overline{U}_i \right)$$

It is convenient to think of the elements of $\Delta$ as “properties”. So $A(S)$ is the probability that all of the properties in $S$ hold, and $B(S)$ is the probability that none of the properties in $S$ hold.
Symmetric Inclusion-Exclusion Theorem

\[ A(S) = \sum_{T \subseteq S} (-1)^{|T|} B(T). \]

\[ B(S) = \sum_{T \subseteq S} (-1)^{|T|} A(T). \]

By symmetry, it’s enough to prove the second formula. This follows almost immediately from the right form of inclusion-exclusion; however we will derive it from asymmetric inclusion-exclusion as stated earlier.
Proof. Let us fix $S$. We define two functions on subsets of $S$. We define

$$g(X) = A(S - X) = P\left( \bigcap_{i \in S - X} U_i \right)$$

$$f(X) = P\left( \left( \bigcap_{i \in S - X} U_i \right) \cap \left( \bigcap_{i \in X} \overline{U}_i \right) \right).$$

So $g(X)$ is the probability that all the properties in $S - X$ hold, and $f(X)$ is the probability that all the properties in $S - X$ and none of the properties in $X$ hold. Then

$$g(X) = \sum_{Y \subseteq X} f(Y),$$

so by ordinary inclusion-exclusion,

$$f(X) = \sum_{Y \subseteq X} (-1)^{|X| - |Y|} g(Y).$$
Now let’s take $X = S$ in the last formula. By definition, 
$f(S) = P \left( \bigcap_{i \in S} \overline{U_i} \right) = B(S)$, so

$$B(S) = \sum_{Y \subseteq S} (-1)^{|S| - |Y|} A(S - Y),$$

and replacing $Y$ with $S - T$ gives

$$B(S) = \sum_{T \subseteq S} (-1)^{|T|} A(T).$$
Using the symmetric inclusion-exclusion theorem, we can easily construct symmetric inclusion-exclusion function pairs with probabilistic interpretations. In order to construct symmetric inclusion-exclusion sequence pairs we need to find examples where \( P\left(\bigcap_{i \in S} U_i\right) \) depends only on \( |S| \). This is very easy to do in the finite case, but not so easy in the infinite case.
A simple example (Bernoulli trials)

We flip a coin infinitely many times. Each flip comes up heads with probability $p$ and tails with probability $1 - p$. For $i \in \mathbb{P} = \{1, 2, 3, \ldots \}$ we let $U_i$ be the event that the $i$th flip is a head. Then for any finite subset $S \subseteq \mathbb{P}$

$$A(S) = P\left( \bigcap_{i \in S} U_i \right) = p^{|S|}$$

$$B(S) = P\left( \bigcap_{i \in S} \overline{U}_i \right) = (1 - p)^{|S|}$$

Then symmetric inclusion-exclusion gives

$$(1 - p)^n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} p^k$$
The Pólya-Eggenberger urn model

We have an urn that initially contains \( r \) red balls and \( b \) black balls. At each step we choose a ball at random from the urn. (So initially the probability of picking a red ball is \( r/(r + b) \) and the probability of choosing a black ball is \( b/(r + b) \).) We then replace the ball we have picked and add another ball of the same color. We then repeat this procedure forever.
The Pólya-Eggenberger urn model

We have an urn that initially contains $r$ red balls and $b$ black balls. At each step we choose a ball at random from the urn. (So initially the probability of picking a red ball is $r/(r+b)$ and the probability of choosing a black ball is $b/(r+b)$.) We then replace the ball we have picked and add another ball of the same color. We repeat this procedure forever.
Note:
(i) \( r \) and \( b \) don’t really need to be integers, but they should be positive.
(ii) Instead of adding one new ball at each step, we could add \( c \) balls.
(iii) If we take the limit as \( b, r \to \infty \) so that \( r/b \to p \) then we have coin flips.
What is the probability that the first five balls have the colors RRBBR?

\[
\frac{r}{r+b} \cdot \frac{r+1}{r+b+1} \cdot \frac{b}{r+b+2} \cdot \frac{b+1}{r+b+3} \cdot \frac{r+2}{r+b+4}
\]

What is the probability that the first five balls have the colors BRRBR?

\[
\frac{b}{r+b} \cdot \frac{r}{r+b+1} \cdot \frac{r+1}{r+b+2} \cdot \frac{b+1}{r+b+3} \cdot \frac{r+2}{r+b+4}
\]
In general, the probability that the first \( m + n \) balls are any particular sequence of \( m \) red and \( n \) black balls is

\[
\frac{(r)_m (b)_n}{(r+b)_{m+n}},
\]

where \((\alpha)_n = \alpha(\alpha + 1) \cdots (\alpha + n - 1)\).

As a consequence, we can show that for any disjoint sets of integers \( R = \{i_1, i_2, \ldots, i_m\} \) and \( B = \{j_1, j_2, \ldots, j_m\} \), the probability that the balls with numbers in \( R \) are red and the balls with numbers in \( B \) are black depends only on \( m \) and \( n \), and is therefore

\[
\frac{(r)_m (b)_n}{(r+b)_{m+n}}.
\]

Example:

\[
P(R \cdot B) = P(RBB) + P(RRB)
\]

\[
P(RB \cdot) = P(RBB) + P(RBR)
\]
In general, the probability that the first \( m + n \) balls are any particular sequence of \( m \) red and \( n \) black balls is

\[
\frac{(r)_m (b)_n}{(r + b)_{m+n}}
\]

where \((\alpha)_n = \alpha(\alpha + 1) \cdots (\alpha + n - 1)\).

As a consequence, we can show that for any disjoint sets of integers \( R = \{i_1, i_2, \ldots, i_m\} \) and \( B = \{j_1, j_2, \ldots, j_m\} \), the probability that the balls with numbers in \( R \) are red and the balls with numbers in \( B \) are black depends only on \( m \) and \( n \), and is therefore

\[
\frac{(r)_m (b)_n}{(r + b)_{m+n}}
\]

Example: \( P(R \cdot B) = P(RBB) + P(RRB) \)

\( P(RB \cdot ) = P(RBB) + P(RBR) \)
Note: The urn model is equivalent to a lattice path model in the plane, where a particle starts at the origin, and from the point \((i, j)\) it moves right with probability \((r + i)/(r + b + i + j)\) and up with probability \((b + j)/(r + b + i + j)\). Then the probability that any particular path ending at \((m, n)\) has been taken is \((r)^{m}(b)^{n}/(r + b)^{m+n}\).
Now let $U_i$ be the event that ball $i$ is red, so $\overline{U_i}$ is the event that ball $i$ is black. So with our previous notation, for any set $S$ of positive integers, $A(S)$ is the probability that all the balls with numbers in $S$ are red, and

$$A(S) = \frac{(r)_{|S|}}{(r+b)_{|S|}}.$$  

Similarly, $B(S)$ is the probability that ball $i$ is black for all $i$ in $S$, and

$$B(S) = \frac{(b)_{|S|}}{(r+b)_{|S|}}.$$  

So if we take $S$ to be a set of size $n$, symmetric inclusion-exclusion gives

$$\frac{(b)_n}{(b+r)_n} = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{(r)_k}{(b+r)_k}$$

and the same identity with $b$ and $r$ switched.
This identity is a form of the Chu-Vandermonde summation theorem. The corresponding exponential generating function identity

\[
\sum_{n=0}^{\infty} \frac{(b)_n}{(b + r)_n} \frac{x^n}{n!} = e^x \sum_{n=0}^{\infty} (-1)^n \frac{(r)_k}{(b + r)_k} \frac{x^n}{n!}
\]

is the \(1 F_1\) transformation

\[
_{1}F_{1}\left(\begin{array}{c} b \\ b + r \end{array} \mid x \right) = e^x \ _{1}F_{1}\left(\begin{array}{c} r \\ b + r \end{array} \mid -x \right).
\]
Are there any other symmetric inclusion-exclusion sequence pairs with probabilistic interpretations?

If \((a_n)\) and \((a'_n)\) are symmetric inclusion-exclusion sequences then so is \((a_n a'_n)\). So we can multiply the symmetric inclusion-exclusion sequences that we already have.
Are there any other symmetric inclusion-exclusion sequence pairs with probabilistic interpretations?

If \((a_n)\) and \((a'_n)\) are symmetric inclusion-exclusion sequences then so is \((a_na'_n)\). So we can multiply the symmetric inclusion-exclusion sequences that we already have.
We take $m$ urns, where the $j$th urn starts with $r_j$ red balls and $b_j$ black balls. At each step, we choose a ball from each urn, replace it, and add another ball of the same color. What is the probability that if we do this $n$ times, at each step we choose at least one black ball? We let $U_i$ be the event that at the $i$th step all the balls chosen are red. Then for any finite set $S \subseteq P$, $A(S)$ is the probability that for all the steps in $S$, all of the balls chosen are red. If $|S| = n$, this probability is

$$
\frac{(r_1)_n}{(r_1 + b_1)_n} \frac{(r_2)_n}{(r_2 + b_2)_n} \cdots \frac{(r_n)_n}{(r_m + b_m)_n}.
$$

The complementary event $\overline{U}_i$ is the event that at the $i$th step, at least one black ball is chosen. So the probability that in $n$ steps at least one black ball is chosen at each step is

$$
B(S) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{(r_1)_k}{(r_1 + b_1)_k} \frac{(r_2)_k}{(r_2 + b_2)_k} \cdots \frac{(r_n)_k}{(r_m + b_m)_k},
$$

where $n = |S|$.
This sum may be written as the hypergeometric series

\[
m+1 F_m \left( \begin{array}{c} -n, \quad r_1, \quad r_2, \quad \cdots, \quad r_m \\ r_1 + b_1, r_2 + b_2, \cdots, r_m + b_m \end{array} \bigg| 1 \right).
\]

As a corollary, we get that this hypergeometric series is positive, as long as the \( r_i \) and \( b_i \) are positive real numbers.

It is not too difficult to prove this result analytically, using the integral representation

\[
\frac{(r)_k}{(r + b)_k} = \frac{\Gamma(r + b)}{\Gamma(r)\Gamma(b)} \int_0^1 x^{k+r-1}(1 - x)^{b-1} \, dx;
\]

however, the combinatorial proof also shows that the sum can be expressed as a quotient of polynomials with positive coefficients.
This sum may be written as the hypergeometric series

\[ \text{$_{m+1}F_m\left(\begin{array}{c} -n, \ r_1, \ r_2, \cdots, \ r_m \\ r_1 + b_1, \ r_2 + b_2, \cdots, \ r_m + b_m \end{array} \bigg| 1 \right)$} \]

As a corollary, we get that this hypergeometric series is positive, as long as the $r_i$ and $b_i$ are positive real numbers.

It is not too difficult to prove this result analytically, using the integral representation

\[ \frac{(r)_k}{(r + b)_k} = \frac{\Gamma(r + b)}{\Gamma(r)\Gamma(b)} \int_0^1 x^{k+r-1}(1 - x)^{b-1} \, dx; \]

however, the combinatorial proof also shows that the sum can be expressed as a quotient of polynomials with positive coefficients.
There is one nontrivial case in which we can evaluate this sum. If \( m = 2, \ b_1 = n + 1, \) and \( r_1 = b_2 + 2r_2 - 1, \) the hypergeometric series becomes

\[
\mathbf{3F2}\left(\begin{array}{c}
-n, \ b_2 + 2r_2 - 1, \ r_2, \\
n + b_2 + 2r_2, \ b_2 + r_2, \ 1
\end{array}\right),
\]

which can be evaluated by Dixon’s theorem to give

\[
\frac{(b_2 + 2r_2)_n(\frac{1}{2} + \frac{1}{2} b_2)_n}{(\frac{1}{2} + \frac{1}{2} b_2 + r_2)_n(b_2 + r_2)_n}
\]
There is another special case of this sum that has appeared in the literature. Suppose that \( r_1 = r_2 = \cdots = r_m = r \) and \( b_1 = b_2 = \cdots = b_m = 1 \). Then since

\[
\frac{(r)_k}{(r+1)_k} = \frac{r}{r+k},
\]

the sum is

\[
U_{m,n}(r) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \left( \frac{r}{r+k} \right)^m.
\]

One of the interesting properties of \( U_m(r) \) is the generating function

\[
\sum_{m=0}^{\infty} U_{m,n}(r) z^m = \frac{n!}{(r+1)^n (1-z) \left( 1 - \frac{r}{r+1} z \right) \left( 1 - \frac{r}{r+2} z \right) \cdots \left( 1 - \frac{r}{r+n} z \right)} z,
\]

which is easily verified by partial fraction expansion.
We can make the right-hand side look nicer by rewriting the formula as

\[
\sum_{m=0}^{\infty} \binom{r+n}{n} U_{m,n}(r) \left(\frac{z}{r}\right)^m = \frac{z}{\prod_{i=r}^{r+n} (1 - z/i)}
\]

These numbers up to a constant factor (or special cases) have been considered by

F. N. David and D. Barton
L. M. Smiley
L. Laforest, Ph. Flajolet et al. (analysis of quadtrees)
D. Foata, G.-N. Han, and B. Lass (coupon collector’s problem)
We can make the right-hand side look nicer by rewriting the formula as

\[
\sum_{m=0}^{\infty} \binom{r + n}{n} U_{m,n}(r) \left( \frac{z}{r} \right)^m = \frac{z}{\prod_{i=r}^{r+n} (1 - z/i)}
\]

These numbers up to a constant factor (or special cases) have been considered by

- F. N. David and D. Barton
- L. M. Smiley
- L. Laforest, Ph. Flajolet et al. (analysis of quadtrees)
- D. Foata, G.-N. Han, and B. Lass (coupon collector’s problem)
Conclusion

We have found a framework for symmetric inclusion-exclusion function pairs, and given some interesting examples of symmetric inclusion-exclusion sequence pairs. But we do not have a general method for constructing symmetric inclusion-exclusion sequence pairs with probabilistic interpretations.