Secret Santa, Generalized

(and Some Enumerative Problems)

Rob Donnelly

May 17, 2006

The Original Secret Santa Problem

The Original Secret Santa Problem Suppose n people attend a party. Each guest brings a gift. The gifts are placed in a bin. Each guest blindly picks one gift from the bin. What is the probability that no guest takes home the gift that he/she brought to the party?

Example If there are 9 guests at the party, a way to depict one possible scenario is as a permutation on the set $\{1, 2, \ldots, 9\}$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 6 & 4 & 5 & 9 & 7 & 8 & 3 \end{pmatrix} \begin{array}{l} \longleftarrow \text{ Person} \\ \longleftarrow \text{ Gift he/she takes} \end{array}$$

A mathematical interpretation of Secret Santa The question essentially asks for the proportion of permutations on the set $\{1, 2, \ldots, n\}$ that are *derangements*, i.e. permutations with no fixed points. (The permutation in our example has six fixed points.)

Set-up for our solution

$$\begin{array}{rrcl} & \mathcal{S}_n & \overset{(def)}{:=} & \text{the permuations on } \{1, \ldots, n\} \\ \text{Then} & |\mathcal{S}_n| & = & n! \end{array}$$

$$\begin{array}{rrcl} & \mathcal{D}_n & \overset{(def)}{:=} & \text{the set of all derangements on } \{1, \ldots, n\} \\ \text{Then} & \dfrac{|\mathcal{D}_n|}{n!} & = & \text{proportion of permutations on } \{1, \ldots, \} \text{ that are derangements} \end{array}$$

$$Fix(p_1, p_2, \ldots, p_k) \overset{(def)}{:=} \text{the set of permutations in } \mathcal{S}_n \text{ that fix } p_1, p_2, \ldots, p_k$$

IMPORTANT NOTE: $Fix(2, 4, 5, 7, 8) \subset Fix(2, 4, 7) \subset Fix(2, 7)$

A "sloppy" way to count $|\mathcal{D}_n|$

$$(*) \qquad |\mathcal{D}_n| = |\mathcal{S}_n| - \begin{pmatrix} |Fix(1)| \\ + \\ |Fix(2)| \\ + \\ \vdots \\ + \\ |Fix(n)| \end{pmatrix} + \begin{pmatrix} |Fix(1,2)| \\ + \\ |Fix(1,3)| \\ + \\ \vdots \\ + \\ |Fix(n-1,n)| \end{pmatrix} - \begin{pmatrix} |Fix(1,2,3)| \\ + \\ |Fix(1,2,4)| \\ + \\ \vdots \\ + \\ |Fix(n-2,n-1,n)| \end{pmatrix} + \cdots$$

The assertion of this identity is justified by the Claim below. Apparently many fixed point permutations will be thrown out multiple times and added back multiple times to the count on the right hand side.

Claim *On the right hand side of* $(*)$*, the net result is that each fixed point permutation is thrown out exactly once.*

Original Secret Santa, continued

Proof of Claim. Consider a permutation σ which has exactly k fixed points.

<u>Example</u> The permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 6 & 4 & 5 & 9 & 7 & 8 & 3 \end{pmatrix}$ has $k = 6$ fixed points

$$\begin{array}{llrr}
\text{Thrown out } \binom{6}{1} & \text{times:} & -\binom{6}{1} & -6 \\
\text{Added back } \binom{6}{2} & \text{times:} & +\binom{6}{2} & +15 \\
\text{Thrown out } \binom{6}{3} & \text{times:} & -\binom{6}{3} & -20 \\
\text{Added back } \binom{6}{4} & \text{times:} & +\binom{6}{4} & +15 \\
\text{Thrown out } \binom{6}{5} & \text{times:} & -\binom{6}{5} & -6 \\
\text{Added back } \binom{6}{6} & \text{times:} & +\binom{6}{6} & \underline{+1} \\
 & & & -1
\end{array}$$

In general, here's how often σ is thrown out and added back in our calculation $(*)$:

$$-\binom{k}{1} + \binom{k}{2} - \binom{k}{3} + \cdots + (-1)^k \binom{k}{k}$$

But, we have the following identity involving alternating sums of binomial coefficients (why?):

$$\binom{k}{0} - \binom{k}{1} + \binom{k}{2} - \binom{k}{3} + \cdots + (-1)^k \binom{k}{k} \;=\; 0$$

And so

$$-\binom{k}{1} + \binom{k}{2} - \binom{k}{3} + \cdots + (-1)^k \binom{k}{k} \;=\; -1$$

$\square$

Conclusion: The proportion of derangements is (approximately) $\frac{1}{e}$ (!!!)

$$\begin{aligned}
|\mathcal{D}_n| &= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \binom{n}{3}(n-3)! + \cdots + (-1)^n \binom{n}{n}(n-n)! \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k \, (n-k)! \\
\frac{|\mathcal{D}_n|}{n!} &= \sum_{k=0}^{n} \frac{1}{n!} \, \frac{n!}{k!\,(n-k)!} (-1)^k \, (n-k)! \\
&= \sum_{k=0}^{n} \frac{(-1)^k}{k!} \\
&= 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \\
&\approx \frac{1}{e}
\end{aligned}$$

A Generalized Secret Santa Problem

Motivating our generalization Recently Wayne Bell passed along to me the following question of Mark Galloway: Shuffle a deck of 52 cards, and then lay them down in a row. What are the chances that you will have an "Ace" among the first four cards, a "2" among the next four cards, etc, or a "King" among the last four cards?*

It will be easier to address the negative version of this question: What are the chances that you not will have an "Ace" among the first four cards, a "2" among the next four cards, etc, or a "King" among the last four cards?

A Mathematical Reformulation Fix positive integers a and b and set $n := ab$. In what follows, a will play the role of the suits (so $a = 4$ in the motivating question) and b will play the role of the denominations of the cards (so $b = 13$ in the motivating question).

$\boxed{a\text{-block}}$ For $1 \leq k \leq b$, the kth *a-block* of $\{1, \ldots, n\}$ is the set $B_k := \{(k-1)a + 1, .., ka\}$.

$\boxed{a\text{-derangement}}$ An *a-derangement* of $\{1, \ldots, n\}$ is a permutation σ in $\mathcal{S}_n$ such that for all $i \in \{1, \ldots, n\}$, the element $\sigma(i)$ is not in the same a-block of $\{1, \ldots, n\}$ as the element i.

$\boxed{\mathcal{D}_{n,a}}$ We let $\mathcal{D}_{n,a}$ denote the set of all a-derangements of $\{1, \ldots, n\}$.

$\boxed{a\text{-fixed element}}$ If $\sigma(i)$ is in the the same a-block as i, then we say i is an *a-fixed element* for the permutation σ, and we say σ *a-fixes* the element i.

$\boxed{Fix_a(p_1, \ldots, p_r)}$ We let $Fix_a(p_1, \ldots, p_r)$ denote the set of all permutations in $\mathcal{S}_n$ that a-fix elements $p_1, \ldots, p_r$.

So we seek to count the number $|\mathcal{D}_{n,a}|$ of a-derangements of $\{1, \ldots, n\}$ and also determine what proportion of all permutations on the set $\{1, 2, \ldots, n\}$ are a-derangements.

As before, a "sloppy" way to count $|\mathcal{D}_{n,a}|$

$$(**) \qquad |\mathcal{D}_{n,a}| == |\mathcal{S}_n| - \begin{pmatrix} |Fix_a(1)| \\ + \\ \\ |Fix_a(2)| \\ + \\ \vdots \\ + \\ |Fix_a(n)| \end{pmatrix} + \begin{pmatrix} |Fix_a(1,2)| \\ + \\ \\ |Fix_a(1,3)| \\ + \\ \vdots \\ + \\ |Fix_a(n-1,n)| \end{pmatrix} - \begin{pmatrix} |Fix_a(1,2,3)| \\ + \\ \\ |Fix_a(1,2,4)| \\ + \\ \vdots \\ + \\ |Fix_a(n-2,n-1,n)| \end{pmatrix} + \cdots$$

Proof. As before, on the right hand side of $(**)$, the net result is that each a-fixed point permutation is thrown out exactly once. $\square$

* Actually Wayne and Mark asked about having an "Ace" in the 1st, 14th, 27th, or 40th position, or a "2" in the 2nd, 15th, 28th, or 41st position, etc. But it is clear that there is a one-to-one correspondence between such permutations and permutations satisfying this requirement of an "Ace" among the first four cards etc.

Generalized Secret Santa, continued

What do we do now? So far the set-up is the same as before. In working with the right-hand side of (**), it is clear that we need to understand the quantity

$$\sum_{\substack{\text{subsets } \{p_1,\ldots,p_k\} \\ \text{of } \{1,\ldots,n\}}} |Fix_a(p_1,\ldots,p_k)|$$

Compositions A *composition* of a positive integer k is a sequence of positive integers whose sum is k. If $\mathbf{c} = (c_1,\ldots,c_i)$ is a composition of k, then we say $\mathbf{c}$ has i *parts.* We let $\mathcal{C}(k,i,a)$ denote the compositions of k with i parts and with each part no larger than a.

In what follows, the quantity $(s)_t := s(s-1)(s-2)\cdots(s-t+1)$ when t is a positive integer.

Proposition $$\sum_{\substack{\text{subsets } \{p_1,\ldots,p_k\} \\ \text{of } \{1,\ldots,n\}}} |Fix_a(p_1,\ldots,p_k)| = \sum_{i=1}^{k} \sum_{(c_1,\ldots,c_i)\in\mathcal{C}(k,i,a)} \binom{b}{i} \frac{(a)_{c_1}^2}{c_1!}\cdots\frac{(a)_{c_i}^2}{c_i!}(n-k)!$$

We'll prove this in a moment. For brevity, let $R_k := \sum_{i=1}^{k} \sum_{(c_1,\ldots,c_i)\in\mathcal{C}(k,i,a)} \binom{b}{i} \frac{(a)_{c_1}^2}{c_1!}\cdots\frac{(a)_{c_i}^2}{c_i!}$. (When $k=0$ we take $R_0 = 1$ since it is an empty sum.) Then by (**) and this proposition, we have:

$$\begin{aligned} |\mathcal{D}_{n,a}| &= \sum_{k=0}^{n}(-1)^k R_k\,(n-k)! \\ \frac{|\mathcal{D}_{n,a}|}{n!} &= \sum_{k=0}^{n}(-1)^k \frac{R_k}{(n)_k} \end{aligned}$$

So, we have an answer to our Generalized Secret Santa problem provided we can reasonably work with the numbers R_k. Notice that when $a=1$ we're back in the Original Secret Santa setting. In this case, $b=n$ and $\mathcal{C}(k,i,1)=\emptyset$ unless $i=k$, in which case $\mathcal{C}(k,k,1) = \{(1,1,\ldots,1)\}$. Then $R_k = \binom{n}{k}$.

Proof of the proposition We group the subsets $\{p_1,\ldots,p_k\}$ according to their "composition type." By this we mean: Let i be the number of different a-blocks containing the elements $\{p_1,\ldots,p_k\}$, which we will denote $B_{r_1},\ldots,B_{r_i}$. For $1\le j\le i$, let

$$c_j := |B_{r_j}\cap\{p_1,\ldots,p_k\}|,$$

the number of elements from $\{p_1,\ldots,p_k\}$ that are in a-block B_{r_j}. Notice that $\mathbf{c} := (c_1,\ldots,c_i)$ is a composition of k and that each part is no larger than a. We will say $\{p_1,\ldots,p_k\}$ has *composition type* $\mathbf{c}$.

Generalized Secret Santa, continued

Proof, continued If we fix a composition $\mathbf{c} = (c_1, \ldots, c_i)$ in $\mathcal{C}(k, i, a)$ and on the left-hand side of the following identity sum over all subsets $\{p_1, \ldots, p_k\}$ of composition type $\mathbf{c}$ then

$$\sum |Fix_a(p_1, \ldots, p_k)| = \binom{b}{i}\binom{a}{c_1}(a)_{c_1}\binom{a}{c_2}(a)_{c_2}\cdots\binom{a}{c_i}(a)_{c_i}\,(n-k)!$$

To see this, we have $\binom{b}{i}$ counting the number of ways to choose i a-blocks $B_{r_1}, \ldots, B_{r_i}$ from among the blocks $B_1, \ldots, B_b$. To understand the quantity $\binom{a}{c_1}(a)_{c_1}$, we choose c_1 a-fixed elements from among the elements in the a-block B_{r_1} (there are $\binom{a}{c_1}$ such choices), and we have $(a)_{c_1}$ choices for where these c_1 elements get sent by a permutation which a-fixes them. We can similarly understand the quantities $\binom{a}{c_j}(a)_{c_j}$ for all $1 \le j \le i$. Finally there are $(n-k)!$ ways to freely assign the remaining $n-k$ elements of $\{1, \ldots, n\}$. To finish the proof, check that

$$\binom{a}{c_j}(a)_{c_j} = \frac{(a)_{c_j}^2}{c_j!} \qquad \square$$

An interpretation of the numbers R_k The topic of permutations avoiding certain patterns is well-studied, so it was natural at some point to consult *Enumerative Combinatorics, Vol. I* by the eminent Richard Stanley. There I was reminded of something called "Rook Theory" which leads to the following interpretation of the numbers R_k.

Color b square $a \times a$ regions along the "main diagonal" of an $n \times n$ chess board. Then R_k (for $0 \le k \le n$) counts the number of ways of placing k non-attacking rooks on the colored regions. The proof of this is entirely similar to our proof of the above proposition.

This particular generalization of the Secret Santa Problem was not in Stanley's text, and I wonder what is known about it. Can we take the absence of a nice closed formula for our R_k's in his book as evidence of the non-existence of such a formula?

Answering the original question Using the above formula for R_k, I calculated R_0, R_1, R_2, R_3, R_4, R_5, and R_6 by hand (with the aid of a hand-held calculator). Here's what I got:

R_0	R_1	R_2	R_3	R_4	R_5	R_6
1	208	20,904	1,352,416	63,317,176	2,286,355,968	66,274,500,864

Performing these calculations by hand was time-intensive. One part of the exercise was to generate all compositions of k (for $1 \le k \le 6$) with no part larger than $a = 4$. Also, the numbers seemed to be getting prohibitively large fairly quickly, which made me pessimistic about being able to follow through with the calculations even if they were automated in MAPLE. It is also instructive to use partial sums $\sum_{k=0}^{K}(-1)^k \frac{R_k}{(n)_k}$ to approximate $\frac{|\mathcal{D}_{n,a}|}{n!} = \sum_{k=0}^{n}(-1)^k \frac{R_k}{(n)_k}$. I did this by hand for $K = 0$, 1, 2, 3, 4, 5, and 6. It was clear that this was not converging very quickly to the desired probability:

K	0	1	2	3	4	5	6
Partial sum	1	-3	4.882353	-5.316863	4.428139	-2.902857	-0.731986

Generalized Secret Santa, continued

Answering the original question, continued At this point I wrote a MAPLE program that would take as input n, a, b, and an upper limit value K for the approximating partial sum. I checked the veractiy of the program by confirming the above data from the calculations I had performed by hand and also by making sure the program returned the correct result for the case we know: the Original Secret Santa Problem where $a = 1$. Everything checked out OK. Here's what the program returned for $n = 52$, $a = 4$, $b = 13$, and K values from 0 through 19:

K	0	1	2	3	4	5	6
Time	—	—	—	—	—	—	—
Memory	—	—	—	—	—	—	—
Partial sum	1	-3	4.882353	-5.316863	4.428139	-2.902857	1.618489

K	7	8	9	10	11	12	13
Time	—	—	—	—	1 sec	2 sec	5 sec
Memory	—	—	—	—	—	—	—
Partial sum	-0.731986	0.319012	-0.091439	0.050252	0.006599	0.018696	0.015661

K	14	15	16	17	18	19
Time	10 sec	31 sec	112 sec	494 sec	2551 sec	NA
Memory	2 meg	3.25 meg	6.75 meg	18.6 meg	50.8 meg	Out of memory
Partial sum	0.01635395	0.01620917	0.01623693	0.01623203	0.01623283	—

It appears that approximately 1.623% is the probability of dealing 52 cards in a row without an "Ace" among the first four cards, a "2" among the next four cards, etc.

We devised a MAPLE simulation to shuffle and deal 10,000 times, i.e. generate 10,000 random permutations of the set $\{1, 2, \ldots, 52\}$. Of these, 156, or 1.56%, were 4-derangements of $\{1, 2, \ldots, 52\}$.

Questions about this approach

First, is an explicit closed formula for the "rook numbers" R_k available?

Second, if not, is there a more efficient way to calculate these numbers (for example, by a more straightforward recurrence than the one I use in my program)?

Third, in what sense do the partial sums "converge" to the correct probability? What can be said concretely about the errors in the estimating partial sums?

Fourth, can we say whether the absolute values of the terms of the partial sum are eventually decreasing?

Fifth, what about asymptotic considerations? What happens as $b \to \infty$? (When $a = 1$, then as $b \to \infty$ we have the probability converging to $\frac{1}{e}$.) Is there a limit? Is the limit an irrational number? A transcendental number?