

Derangements

Brian Conrey and Tom Davis
conrey@aimath.org and tomrdavis@earthlink.net
<http://www.geometer.org/mathcircles>
March 23, 2000

1 Seating Mixup

Imagine that Yankee Stadium is completely sold out, but when the ticket holders arrive, they choose seats entirely at random. What is the probability that at least one person is seated in the seat indicated by his ticket?

Obviously, you need to know how many seats are there—if there were only one seat in the stadium, then that person has to be in his seat, so the probability that he is correctly seated is 1. For the purposes of this problem, assume that there are 50000 seats in Yankee Stadium.

1.1 Experiments

It will turn out that there is not too much difference in the answer if there are 20 seats or 50000 seats, but that's not obvious at first. One nice way to approach the problem is to do a few experiments that require only a deck of playing cards.

Let's run some experiments with a mini Yankee Stadium that contains only 13 seats. Select all of the spades and all of the hearts from the deck. Now place all the spades, in order, in a line like this:

A♠ 2♠ 3♠ 4♠ 5♠ 6♠ 7♠ 8♠ 9♠ 10♠ J♠ Q♠ K♠

Now shuffle the hearts and deal them out underneath the spades, making a pattern that looks like this:

A♠ 2♠ 3♠ 4♠ 5♠ 6♠ 7♠ 8♠ 9♠ 10♠ J♠ Q♠ K♠
7♥ 5♥ A♥ 9♥ 8♥ Q♥ 2♥ 3♥ J♥ 10♥ K♥ 4♥ 6♥

For the particular deal above, the 10 of hearts lands under the 10 of spades, so there is a match—the person with the ticket number 10 wound up in seat number 10.

Repeat this experiment many times. Shuffle the hearts and repeatedly deal them out under the spades. After each deal, check to see if there is a match, and keep track of the number of times there is at least one match and the number of times there is no match. From this data, calculate an approximate probability, based on the experiment.

Now change the experiment and do the same thing, but with only 8 cards from each suit (ace through 8 of hearts and spades). Calculate an approximate probability based on this smaller deck.

If you can program computers, you might try to simulate this experiment for thousands of deals. A sample program written by the author generated the following results with 100000 deals using decks of cards of size 1, 2, 3, 4, 5, 10, and 30. The number in the left column is the number of cards in the deck; the number in the right column is the number of deals (out of 100000) where at least one of the cards was in the correct spot. The code for this program, written in C, appears at the end of this document.

1	100000
2	49735
3	66763
4	62655
5	63341
10	63170
30	63059

This makes some sense—if there’s only one card, there will always be a match, so there have to be 100000 matches. If there are two cards, there are only two possible orderings. One of those orderings switches the cards and the other doesn’t, so half the deals should match. The experimental result of 49735 out of 100000 is very close to 50%.

2 Calculations for Small Decks

Let’s see if we can work out the exact results, at least for a few small decks. We can check our calculations against the experimental results above. We have already done this for decks of size 1 and 2.

For a deck of size 3, there are 6 possible deals of the cards, all equally likely. Let’s just call the cards 1, 2, and 3. Here are the possible deals, with a ● mark in the final column if there’s a match:

1	2	3	
1	2	3	●
1	3	2	●
2	1	3	●
2	3	1	
3	1	2	
3	2	1	●

Four of the six deals have a match, so we expect that roughly $2/3 = .666\dots$ of the hands should have a match. Our experimental result of $66763/100000 = .66763$ is pretty close.

Now let’s do the same thing for a four card deck:

1	2	3	4	
1	2	3	4	•
1	2	4	3	•
1	3	2	4	•
1	3	4	2	•
1	4	2	3	•
1	4	3	2	•
2	1	3	4	•
2	1	4	3	
2	3	1	4	•
2	3	4	1	
2	4	1	3	
2	4	3	1	•
3	1	2	4	•
3	1	4	2	
3	2	1	4	•
3	2	4	1	•
3	4	1	2	
3	4	2	1	
4	1	2	3	
4	1	3	2	•
4	2	1	3	•
4	2	3	1	•
4	3	1	2	
4	3	2	1	

In this example, there are 15 cases out of the 24 possible deals where there is a match. The probability should be $15/24 = .625$, again not too far from our experimental result of $62655/100000 = .62655$.

To do an exhaustive list for the case of a 5 card deck would require 120 deals which we could do, but it would not be pleasant. But look at the case above—of course if there's a 1 in the first slot, there will be a match, so all 6 deals like that will be matches. But if there is a 2 dealt first, there are 3 matches. Similarly, a 3 dealt first or a 4 dealt first yield exactly the same number (3) of matches. Therefore there are $6 + 3 \cdot 3 = 15$ deals with at least one match.

For the case of a 5 card deck, there are 5 possible cards that can come first. If a 1 comes first, all have a match, for 24 combinations. All four other possibilities for the first card are similar, so we merely need to count them for a 2 in the first slot. Do this as an exercise, and you'll find that there are 13, so the grand number of deals in which there is at least one match is $24 + 4 \cdot 13 = 76$. So 76 out of 120 deals gives a probability of a match of $76/120 = .6333 \dots$, again close to the experimental result of $63341/100000 = .63341$.

It is possible to count larger decks, and if you're interested, it is an interesting experience, but let's see what we have so far:

The table below has four columns: the size of the deck, the number of deals with at

least one match, the total number of possible deals, and the probability of at least one match:

1	1	1	$1/1 = 1$
2	1	2	$1/2 = .5$
3	4	6	$4/6 = .666\dots$
4	15	24	$15/24 = .625$
5	76	120	$76/120 = .6333\dots$

Notice that the probabilities seem to alternate, going down, then up, then down, but also seeming to converge.

In fact, look at the differences between the probabilities: $1 - 1/2 = 1/2$, $1/2 - 4/6 = -1/6$, $4/6 - 15/24 = 1/24$, and finally, $15/24 - 76/120 = -1/120$.

The signs alternate, but the fractions are all of the form $\pm 1/n!$, where $n! = n(n - 1)(n - 2) \cdots 1$.

So just looking at the data we have, it appears that the probability of getting at least one match from a random deal of a deck of n cards is:

$$\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots \pm \frac{1}{n!}.$$

3 Counting Derangements

Now let's take a look at a related problem. Imagine that you have a bunch of people who list their food preferences. For the first example, imagine that we just ask everybody about two foods: cake and peanuts. Suppose 10 people don't like either, 5 people like cake but not peanuts, 11 people like peanuts and not cake, and 14 people like both. So there are $10 + 5 + 11 + 14 = 40$ total people.

Another way of looking at the people is this. Of the 40 people, 19 people like cake (the 5 who like cake only and the 14 who like both cake and peanuts), and 25 people like peanuts. Suppose we know these numbers and that there are 40 total people and we try to work out how many people don't like either. The first approach is to start with the 40 and subtract off the people who like one or the other: $40 - 19 - 25 = -4$. Clearly this isn't the right answer—we can't have a negative number of people but what went wrong?

The answer is that we subtracted off the people who like both cake and peanuts more than once—one time when we subtracted the cake people and one time when we subtracted the peanut people. We only wanted to subtract the cake-and-peanut people once, but we subtracted them twice, so if we add them in, we'll get the right result: $40 - 19 - 25 + 14 = 10$, and 10 is the correct number of people who like neither.

Now consider a more complex example, where there are three possible foods: cake, peanuts, and turnips. Every person either likes or dislikes each, so there are eight different categories of people. In the chart below, all the possibilities are listed, together with the number of people who fall into each category:

C	P	T	#
			11
		•	5
	•		8
	•	•	2
•			20
•		•	15
•	•		4
•	•	•	7

If there is a bullet (•) in a column, that means the person likes that particular food. “C”, “P”, and “T” stand for “cake”, “peanuts”, and “turnips”. The fourth line indicates that there are 2 people who like peanuts and turnips, but do not like cake, et cetera.

Every possible set of likes and dislikes is listed, and there are 72 total people. Of these people, 46 like cake (and perhaps other things). Similarly, 21 like peanuts, and 29 like turnips.

If we try our defective method for counting the number of people who don’t like any of the foods, the first approximation is $72 - 46 - 21 - 29 = -24$. But we know what’s wrong—many people were subtracted more than once.

There are 11 people who like cake and peanuts (and may or may not like turnips), 22 people who like cake and turnips, and 9 people who like peanuts and turnips, so maybe we’d better add them back in: $72 - 46 - 21 - 29 + 11 + 22 + 9 = 18$. This is still wrong. The reason is that 7 people like all three. They appear in the original list of 72, but were subtracted out 3 times originally. But then they were added in 3 times when we added the people who like at least two foods. We need to subtract them out again, and when we take the 7 from 18, we get 11—exactly the number who don’t like anything.

This same general scheme will work for any number of items. To find the number of people who don’t like any of them, start with the total number of people, subtract off all the people who like each of the individual items, then subtract off the people who like at least two, add in the number who like at least 3, and so on.

We can do exactly the same sort of thing with the Yankee Stadium problem. If there are N seats, there are $N!$ possible arrangements. We want to count the number of “derangements”—orderings where *none* of the people get the correct seat. To do this, begin with $N!$ and subtract off the ones where there is a match.

This is an easy example of what we just did—in the previous examples, we had varying numbers of people who liked different combinations of food, but here we just consider a single version of each seating arrangement.

How many arrangements are there with person 1 in seat 1? Well, the rest of the $N - 1$ people can be arranged in any of $(N - 1)!$ ways, so there are $(N - 1)!$ such arrangements. There are a similar number of arrangements with person 2 in seat 2, et cetera. So we’ll subtract all those off. But then we’ve subtracted many arrangements more than once. Some of the arrangements have the correct person in both seats 1 and 2, and we subtracted that arrangement twice.

How many arrangements have person 1 and 2 in the correct seats? Well, there are $N - 2$ remaining seats, and those can be arranged in any of $(N - 2)!$ ways. We need to add in all of those, then subtract off all the cases where there are at least 3 people in the right seats, then add in the cases with 4, et cetera.

So there are $\binom{N}{1}$ ways to pick one person with the correct seat, $\binom{N}{2}$ ways to pick two people with the correct seat, and so on.

Thus, the total number of derangements is this:

$$\begin{aligned} & N! - \binom{N}{1}(N-1)! + \binom{N}{2}(N-2)! - \binom{N}{3}(N-3)! + \dots \\ &= N! - \frac{N(N-1)!}{1!} + \frac{N(N-1)(N-2)!}{2!} - \frac{N(N-1)(N-2)(N-3)!}{3!} + \dots \\ &= N! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots \right). \end{aligned}$$

But there are $N!$ rearrangements, so to get the probability of a derangement, we have to divide by $N!$, giving the following probability for a derangement:

$$\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots \right).$$

In the previous section we were counting the opposite of derangements—we were counting cases where at least one slot was correctly filled, so we need to subtract the number above from 1 to get that result, and thus we obtain the guess we made previously:

$$\left(\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \dots \right).$$

For reasonably large N , the numerical value of this is roughly .632120558829.

If you happen to know the Taylor series for e^x , you can get a quick approximation for large N :

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$

So

$$e^{-1} = \frac{1}{e} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots$$

From this it's easy to derive the result we want. The probability of at least one matching seat is approximately given by:

$$1 - \frac{1}{e} = .632120558829\dots$$

4 Card Deck Simulation Code

```
#include <stdio.h>

#define DEALS 100000

main(int argc, char **argv)
{
    int cards[1000];
    int count;
    int i, j, c1, c2, shuffle, tmp, hit;
    int success = 0;

    if (argc != 2) {
        fprintf(stderr, "Usage: %s <number of cards>\n", argv[0]);
        exit(1);
    }
    count = atoi(argv[1]);
    if (count <= 0) count = 10;
    for (i = 0; i < count; i++) cards[i] = i;
    for (i = 0; i < DEALS; i++) {
        shuffle = 30 + rand()%1001;
        for (j = 0; j < shuffle; j++) {
            c1 = rand()%count;
            c2 = rand()%count;
            tmp = cards[c1];
            cards[c1] = cards[c2];
            cards[c2] = tmp;
        }
        hit = 0;
        for (j = 0; j < count; j++)
            if (cards[j] == j) hit = 1;
        if (hit) success++;
    }
    printf("Deals: %d; Successes: %d\n", DEALS, success);
}
```