International Journal of Pure and Applied Mathematics
Volume 81 No. 6 2012, 821-830
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu
ijpam.eu

UNEXPECTED PROBABILITIES

Yutaka Nishiyama

Department of Business Information
Faculty of Information Management
Osaka University of Economics
2, Osumi Higashiyodogawa Osaka, 533-8533, JAPAN

Abstract: This article presents 2 unexpected probabilities. If there are at least 23 people in a class, then the probability that there is a shared birthday exceeds one half. This is explained by the idea of complementary events. What are the chances of there being a student who ends up in the same seat when everyone in the class is reseated? This probability is around $\frac{2}{3}$ irrespective of the number of students, and was solved by Montmort (1708).

AMS Subject Classification: 60A02, 00A09, 97A20
Key Words: likelihood of shared birthdays, complementary events, probability of ending up in the same seat, Montmort's theory, Venn diagram

1. The Likelihood of Shared Birthdays

There are all kinds of problems in probability, and the study of probability gets more and more interesting when there is a big gulf between expectation and reality. I refer to examples of this gulf between expectation and reality as 'unexpected probabilities', and in this chapter I'd like to introduce two such probability problems.

What kind of value do you expect for the probability that two people in a class share the same birthday? There are about 40 people in one of my classes. I

Received: August 18, 2012

ask students to estimate the probability that two people have the same birthday, assuming there are 365 days in one year. Most students estimate on the side of 'no one'. This is because they reason that while there are 40 students in the class, there are 365 days in the year, so the chance of a shared birthday is $\frac{40}{365}$. They also suppose that the first time there would be students with the same birthday is with a class of 366 students.

One's understanding of probabilities cannot be deepened without trying out some practical experiments. I take up this example every year in my lectures on information mathematics intended for students in the humanities. I have students say their birthdays, and after hearing from all the students once, I ask students whose birthdays were duplicated to raise their hands. Enough students raise their hands at this point that it's safe to say there's always some. Owing to the fact that they estimated the probability as being around $\frac{40}{365}$, students find this considerable discrepancy mysterious. It is at this point that I provide a mathematical explanation of probability.

This is a problem which is solved using the concept of complementary events. The number of people in the class is taken to be n, and the number of days in one year is taken to be 365. First we obtain the probability that everyone in the class has a different birthday. There are 365 days from which the first person's birthday may be chosen, so the probability associated with the first person is $\frac{365}{365}$. There is then one less day in the year from which the second person's birthday may be chosen, leaving 364 days. The probability is $\frac{364}{365}$. For the nth person, the probability associated with their birthday is $\frac{365-(n-1)}{365}$, so for a total of n people, the probability that they all have different birthdays is

$$P_1 = \frac{365}{365} \times \frac{365-1}{365} \times \cdots \times \frac{365-(n-1)}{365}.$$

The probability that there is at least one shared birthday can be obtained as the probability of a complementary event as described above,

$$P = 1 - P_1 = 1 - \frac{365}{365} \times \frac{365-1}{365} \times \cdots \times \frac{365-(n-1)}{365}.$$

Furthermore, when $n = 23$, we have $P = 0.507$. This means that if there are at least 23 people in the class then the probability that there is a shared birthday exceeds one half. Considering this value of 23, it is less than one tenth of the 365 days.

The concept of complementary events is extremely useful. The concepts in the phrases 'denying that all the birthdays are different', and 'at least one birthday is repeated' are expressed using the notions of 'not always' and 'at least', respectively. Problems involving complementary events often come up, so it is surely important to gain familiarity with them. This kind of surprising probability problem is not widely known, and the reason it does not often appear in exams likely stems from the fact that it is difficult to calculate P_1 and P on paper. Despite being an interesting example, forgetting about it just because it is not suited to examinations is not ideal.

Personal computers are now prevalent, and by using spreadsheet software the calculation is rendered straightforward. Calculating the probability of there being a shared birthday, P, while varying the number of people in the class, n, and graphing the results must yield a deeper understanding of probability. When n is small ($n < 10$), P has a small value ($P < 0.1$), but beyond that it rapidly increases so that for $n = 23$ the probability is $P = 0.507$, and for $n > 40$, $P > 0.9$. This birthday-related example can usually be demonstrated practically as a class experiment, so it may be introduced to a class with confidence.

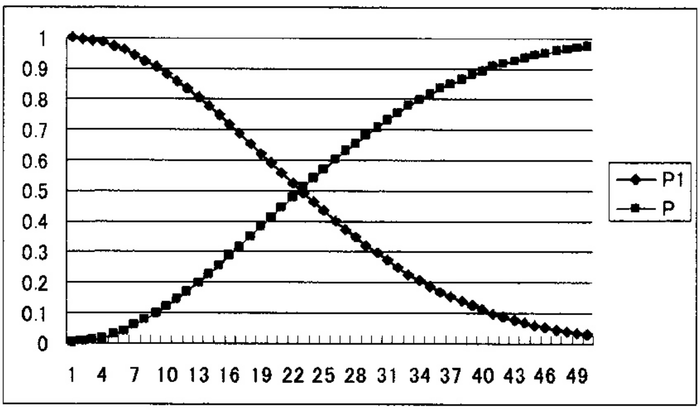

Figure 1: The probability that there is at least one shared birthday in the class

2. What's the Probability that at Least 2 Pairs Share a Birthday?

I lecture about the birthday example every year. In fact, this year when I performed the experiment with a class of 60 people, the result revealed 3 pairs with the same birthday. If there is at least one pair it is sufficient for the lesson, but there were 3 pairs who shared the same birthday. This gave rise to a new further enquiry as to whether the fact that among 60 people there were 3 pairs with the same birthday was an appropriate number. When I performed the experiment with a class 2 years ago, there were 40 students in the class, and there was 1 pair with the same birthday. But this year, for 60 students there were 3 pairs. Moreover, one of the 3 pairs was a pair of fellow students who sat next to each other. Is this an appropriate number after all?

For simulations it is convenient to use random numbers. A confirmation is possible using Visual BASIC, for which the built in random-number generating function RND can be used with a program of about 50 lines. The chance of selecting a pair of people is given by the following formula.

$$ {}_nC_2 = \frac{n!}{(n-2)!2!} $$

For each case, the remaining $(n-2)$ people's birthdays must fall on different days, which has the following probability.

$$ P_2 = {}_nC_2 \times \frac{365}{365} \times \frac{365-1}{365} \times \cdots \times \frac{365-(n-2)}{365} $$

The probability that 2 or more pairs of people share a multiple birthday, or more than 2 people share the same birthday is

$$ 1 - P_1 - P_2. $$

Calculating the equation above for $n = 23$ and $n = 60$ using spreadsheet software reveals that the values take the following form.

n	P_1	$1-P_1$	P_2	$1-P_1-P_2$
23	0.493	0.507	0.363	0.144
60	0.006	0.994	0.034	0.960

Table 1. Probabilities of multiple birthdays ($n = 23, 60$)

When calculating the probability P_2 using spreadsheet software, calculating the numerator and denominator separately as follows causes an overflow during computation.

$numerator = 365 \times {}_nC_2 \times (365-1) \times \cdots \times (365-(n-2))$
$denominator = 365 \times 365 \times 365 \times \cdots \times 365 = 365^{n-1}$

The technique for preventing an overflow during the computation of P_2 is to compute each term completely, in sequence. Proceeding in this way reveals that for 60 people, 3 or 4 pairs with the same birthday is not at all unreasonable.

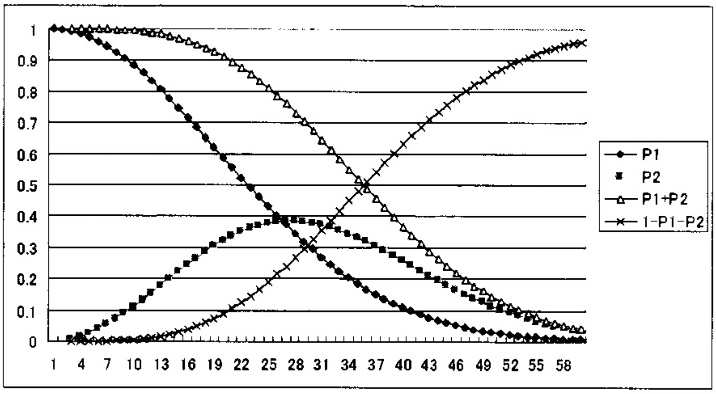

Figure 2: The probabirity that at least 2 pairs have the same birthday

3. The Probability of Ending up in the Same Seat

I'd like to introduce another surprising probability. This is a probability associated with changing seats. What are the chances of there being a student who ends up in the same seat when everyone in the class is reseated? Isn't there someone who has had the unfortunate experience of not getting a different seat when everyone is reseated? Even if it wasn't yourself you must have known a friend who didn't get to change. Let's look and this and explain why it is by no means a great misfortune, but rather an event that occurs with a high probability.

The reseating problem is the same as the 'Secret Santa' Christmas present problem. Should we take the pessimistic perspective that it is a punishment for bad behavior when someone draws the same present that they brought themselves? This is not so, the probability is surprisingly high.

Before writing up a mathematical expression, let's attempt a counting method. Each of the students were denoted by a number 1, 2,c and compared before and after the reseating. Cases where the seat is the same before and after reseating are marked with a circle. When there are 2 students there are 2 seats and the number of ways of arranging them is $2! = 2$. As shown in the first line of Table 2, students 1 and 2 both remain unchanged in 1 case. Since there is 1 case in which a student ends up in the same seat, the probability of this happening is 0.5.

Before reseating	1	2
After reseating	1*	2*
	2	1

Table 2. The case of $n = 2$

In the same way, the case of 3 students is shown in Table 3. There are $3! = 6$ ways of arranging the students, and the cases when a student ends up in the same seat are indicated by circles. For the case in line 1, students 1, 2 and 3 all end up in the same seat. For the case shown in line 2, only student 1 ends up in the same seat. In line 3 only student 3, and in line 6, only student 2 ends up in the same seat. There are 4 cases in which there is a student whose seat does not change, so the probability of this happening is as follows.

$$P = \frac{4}{3!} = \frac{4}{6} = 0.666\cdots$$

Before reseating	1	2	3
	1*	2*	3*
	1*	3	2
After reseating	2	1	3*
	2	3	1
	3	1	2
	3	2*	1

Table 2. The case of $n = 3$

[Question 1] Find the probability that at least one student ends up in the same place for $n = 4$, using a counting method.

There are $4! = 24$ ways of arranging the students after reseating, and there are 15 cases in which there is an unchanged seat. The probability of this happening is therefore

$$P = \frac{15}{4!} = \frac{15}{24} = \frac{5}{8} = 0.625.$$

This counting method is valid, but it's also important to be able to capture the problem mathematically. It is known that for n students, the probability that at least 1 person does not change their seat is

$$P = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots \pm \frac{1}{n!}.$$

Let's think about the derivation of this formula.

When 1 student doesn't change his or her seat, the remaining $(n-1)$ students may be arranged in $(n-1)!$ ways. This works out equivalently whichever of the n students ends up in the same seat, and there are

$$n \times (n-1)! = n!$$

ways in total. In this case some of the remaining students might end up in their original seats, and this situation is discussed below.

In the case that student 1 and student 2 do not change their seats, the remaining $(n-2)$ students may be arranged in $(n-2)!$ ways. There are ${}_nC_2$ ways of choosing the 2 students so there are

$${}_nC_2 \times (n-2)! = \frac{n!}{(n-2)!2!} \times (n-2)! = \frac{n!}{2!}$$

ways in total.

In the same way, the case when 3 students such as student 1, student 2 and student 3 do not change their seats occurs in

$${}_nC_3 \times (n-3)! = \frac{n!}{3!}$$

cases in total.

The relationship above is a little complicated, so for $n = 3$, let's use a Venn diagram to explain the situation (Figure 3). The event that student 1 does not change is denoted A_1, the event that student 2 does not change is denoted A_2, and for student 3, A_3. The event that at least 1 student does not change their seat, A, is thus described as follows.

$$A = A_1 \cup A_2 \cup A_3 = (A_1 + A_2 + A_3) - (A_1A_2 + A_2A_3 + A_3A_1) + A_1A_2A_3$$

For the event A_1, there are 4 cases, *i.e.*, when only student 1 doesn't change seat, when neither student 1 nor student 2 change seats, when neither student 1 nor student 3 change seats, and when neither student 1 nor student 2 nor student 3 changes seats. The equation above reflects the elimination of these overlapping cases. $(A_1 + A_2 + A_3)$ is a simple sum of all the 3 events. The parts that are over-included by duplication are removed by the term $(A_1A_2 + A_2A_3 + A_3A_1)$. However, this again removes too much, so the term $A_1A_2A_3$ is added back in.

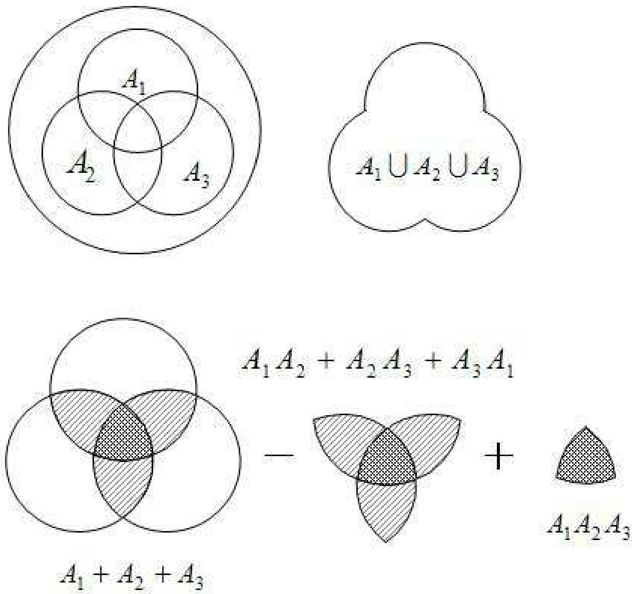

Figure 3: Explanation using a Venn diagram ($n = 3$)

Let's look at the relationship with the $3! = 6$ arrangements for the case $n = 3$ shown in Table 3. $A_1 + A_2 + A_3$ combines a total of 6 instances following reseating corresponding to lines 1 and 2, lines 1 and 6, and lines 1 and 3. $A_1A_2 + A_2A_3 + A_3A_1$ combines a total of 3 instances of line 1. $A_1A_2A_3$ corresponds to a single instance of line 1. This yields

$$A_1 \cup A_2 \cup A_3 = 6 - 3 + 1 = 4$$

arrangements in total. There are also two cases when the seatings change, given by lines 4 and 5.

$$1 - A_1 \cup A_2 \cup A_3 = 2$$

There are 4 arrangements in which at least one seating remains unchanged, and 2 arrangements in which they all change. The total of 6 arrangements can thus be confirmed.

4. Montmort's Theory

In this way, for the general case of n students, event A can be obtained by alternating the signs of the terms and summing. The number of arrangements according to which at least 1 student does not change their seat is therefore given by

$$n! - \frac{n!}{2!} + \frac{n!}{3!} - \frac{n!}{4!} + \cdots \pm 1.$$

Since the total number of arrangements is $n!$ the probability P is

$$P = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots \pm \frac{1}{n!}$$

Table 4 shows the probability P versus the number of students n. It is known that as the value of n increases, the value taken by this formula draws closer to

$$1 - \frac{1}{e} \quad (\approx 0.63212).$$

e is the natural logarithm with value $e \approx 2.71828$. It's amazing that this probability is basically constant, irrespective of the value of n. Moreover, even for infinitely large n, the value of this formula remains larger than 0.6. Even when reseating a class of 1000 students, or when 1000 people exchange Christmas presents, the chances of a student ending up in the same seat, as well as the chances of someone ending up with the present they brought themselves, are over 0.6. This probability differs from that of shared birthdays however, in that it does not tend to 1.

n	P
1	1
2	0.5
3	0.6667
4	0.625
5	0.6333
6	0.6320
7	0.6321

Table 4. Number of students n versus probability P

According to Feller (1968) this problem has a large number of variations leading all the way back to the brilliant solution by Montmort (1708).[1] Two identical sets of n different cards are each arranged in a random order and laid out facing each other. What are the chances of corresponding cards being the same? There are n envelopes and n letters. When a secretary puts the letters in the envelopes indiscriminately, what are the chances of a letter ending up in the right envelope?

The hats deposited in a cloak room are mixed up. Imagine the situation when they are handed back to the guests indiscriminately. When a given person is handed back their own hat, it is considered to be an 'equivalence'. What are the chances of an equivalence occurring? Comparing the chances of an equivalence occurring at an assembly of 8 people and at an assembly of 10000 people, it is surprising to discover that irrespective of the value of n the probability is around $\frac{2}{3}$. The reseating probability thus has many variations, and continues to receive considerable attention.

References

[1] W. Feller, *An Introduction to Probability Theory and its Applications*, 3-rd Edition, John Wiley and Sons (1968).