

# Note – The Birthday Paradox

jacques@ucsd.edu

While this note is not examinable, the combinatorial thinking in this problem is valuable, and instructive to keep in mind. A couple of basic principles come in to play here: the pigeonhole principle, the multiplication principle, and some basic elements of discrete probability.

In class I bet that out of the 42 people in attendance, two would have the same birthday and birth month. This may seem a bit of a foolish bet, since there are 365 possibly days in the year, and only 42 people on the class. If there were more than 365 people in the class, then a win would be assured simply because there must be two people with the same birthday. However, with a bit of combinatorics, we can see that in fact there is a more than 90 percent chance that I bet correctly. This is a counter-intuitive fact which is known more generally as the [birthday paradox](#).

To frame it more mathematically, we have  $n$  people and 365 possible birthdays for each person. Let us count how many ways we can choose  $n$  people all of whose birthdays are [different](#). We do that because, upon division by  $(365)^n$ , we get exactly the probability that all of the  $n$  people have different birthdays. Then subtracting that probability from 1, we get the chance that two do have the same birthday (the [complementary event](#)).

There are 365 choices for the first person's birthday. If the second person's birthday is different, there are only 364 choices for that birthday. After choosing that birthday, there are only 363 choices left. Continuing in this way, we see that there are  $365 - n + 1$  choices for the  $n$ th person's birthday.

So the number of ways of choosing different birthdays for  $n$  people is:

$$365 \cdot 364 \cdot 363 \cdot \dots \cdot (365 - n + 1).$$

A short way of writing this is using [product  \$\Pi\$  notation](#):

$$\prod_{i=1}^n (365 - i + 1).$$

This means multiply out  $365 - i + 1$  for values of  $i$  from 1 to  $n$ , and it is very similar to the [sum  \$\Sigma\$  notation](#) which you will have seen before in calculus. So to conclude this paragraph, the chance that we get  $n$  people all with different birthdays is the above

product divided by  $(365)^n$ , which is

$$\prod_{i=1}^n (365 - i + 1) \cdot \frac{1}{(365)^n} = \prod_{i=1}^n \left(1 - \frac{i-1}{365}\right).$$

Note that we divided each of the  $n$  terms in the first product by 365 to get the second product.

The chance that two people **do** have the same birthday is then

$$1 - \prod_{i=1}^n \left(1 - \frac{i-1}{365}\right).$$

Rather than multiplying the  $n$  terms in the product, it is convenient to show that if  $n$  is big, then the product is small. For example, we claimed that when  $n = 42$ , the product is less than 10 percent. It follows, and we will not do this here, from calculus<sup>1</sup> that the product is actually less than

$$e^{-\frac{1}{365} \sum_{i=1}^n (i-1)} = e^{-\frac{n(n-1)}{730}}.$$

Here we used the fact that

$$\sum_{i=1}^n (i-1) = \frac{1}{2}n(n-1).$$

Putting  $n = 42$ , we get that the chance that two people **do** have the same birthday is greater than

$$1 - e^{-\frac{1722}{730}} \approx 1 - 0.095 \approx 0.905.$$

This confirms that there is more than a 90 percent chance that two out of 42 people have the same birthday. You can check as an exercise that if we have  $n = 100$  people, then the odds become overwhelmingly great that two of the people have the same birthday: the chance that no two have the same birthday is less than one in 750,000!

---

<sup>1</sup>In fact it is true that if  $x_1, x_2, \dots, x_n$  are any non-zero real numbers, then

$$\prod_{i=1}^n (1 - x_i) < e^{-\sum_{i=1}^n x_i}$$

and this follows from the fact that  $1 - x_i < e^{-x_i}$ .