Math Diversion Problem 717

P. Reany

July 13, 2025

With me, everything turns into mathematics.
— Rene Descartes
(P.S. I calculate; therefore I am.)

The following problem is found at:

```
Source: The Ether of Great Mathematical Ideas
Title: The probability of a birthday match in $n$ people
 Presenter: Patrick
```

1 The Problem

A lecture hall is filling up with students for the next lecture. How many students n are needed in the hall so that at least two of them have the same birthday. (Just to clarify: We're looking for the smallest n that satisfies the constraint.)

2 The Preparation

In Diversion Paper 674 and in my monograph of probability theory, we saw that if we partition the sample space into disjoint subsets E_1 and E_2, then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)\,, \tag{1}$$

which can be re-expressed in terms of E and its set complement, $\bar{E}$, as

$$1 = P(E) + P(\bar{E})\,. \tag{2}$$

Now, if we let E be the event that at least two students have the same birthday, then its set complement $\bar{E}$ is the event that no two students have the same birthday. So, by first solving first for $P(\bar{E})$, and then using the formula derived from (2):

$$P(E) = 1 - P(\bar{E})\,, \tag{3}$$

we can arrive at our answer.

3 Solution

Okay, we want that

$$P(E) \geq 0.50\,. \tag{4}$$

Combining this last equation with (3), we have that

$$P(\bar{E}) \leq 0.50\,. \tag{5}$$

So, how should we approach this solution? Let's assign a number n to each student as he or she enters the room. Now, we want to construct an event where there are no birthday duplicates among the n students. After the first two, the probability of getting no duplication is

$$P(\bar{E})\Big|_{n=2} = \left(\frac{365}{365}\right)\left(\frac{364}{365}\right)\,. \tag{6}$$

On the first student there are 365 ways to choose a BD out of 365 possibilities. On the second student there 364 ways to choose a BD out of 365 possibilities, the loss of the BD of the first student ensuring that no duplication occurs up to this point. The third student can have any BD except either of the first two's BDs, giving us

$$P(\bar{E})\Big|_{n=3} = \left(\frac{365}{365}\right)\left(\frac{364}{365}\right)\left(\frac{363}{365}\right)\,. \tag{7}$$

So now we know the pattern:

$$P(\bar{E})\Big|_{n} = \left(\frac{365}{365}\right)\left(\frac{364}{365}\right)\left(\frac{363}{365}\right)\cdots\left(\frac{365-n+1}{365}\right)\,. \tag{8}$$

Now, I have no idea how to easily solve for the smallest n value that will satisfy the constraint in (5). So what I did was to make a successive number of conservative guesses at what n might be, without busting the constraint wildly. So I start with the guess $n = 10$ and got

$$P(\bar{E})\Big|_{n=10} \approx 0.883\,. \tag{9}$$

Then I calculated the next 5 factors to bring n up to 15,

$$P(\bar{E})\Big|_{n=15} \approx 0.883 \times 0.8457\,. \tag{10}$$

Then I calculated the next 5 factors to bring n up to 20,

$$P(\bar{E})\Big|_{n=20} \approx 0.883 \times 0.8457 \times 0.8457 \approx 0.588\,. \tag{11}$$

Then I calculated the next factors, one at a time, until $n = 22$:

$$P(\bar{E})\Big|_{n=22} \approx 0.524\,. \tag{12}$$

At this point, I was pretty sure the next one would do it.

$$P(\bar{E})\Big|_{n=23} \approx 0.524 \times 0.9397 \approx 0.4924\,. \tag{13}$$

If your guess was far off from 23, don't feel bad. Probability theory often produces unintuitive results.