The Birthday Paradox

Sebastian Liu

Problem

Given a set of N individuals, there is some probability that a pair of individuals will have the same birthday. What is the minimum number of individuals such that the probability that a pair of them have the same birthday is greater than 50%? (We are properly ignoring leap years and assuming, for simplicity, that there are 365 days in a year.)

An Intuitive Answer

It should be clear that if $N = 366$, then the probability that there exists at least one pair of individuals sharing a birthday is 100%. A justification of this is left for the reader. It might seem rational, then, to believe that in order to achieve of probability of 50%, we would require $N = \frac{366}{2} = 183$. This number, however, is far too high, as we shall see.

Counting Pairs

We should ask ourselves the following question which is immediately relevant to this problem: given a set of N people, how many unique pairs can be formed? The reason we want to ask this is because given any pair, there is some probability that the two individuals in the pair share a birthday. Formally, we are choosing 2 from a set of N. The general formula for choosing k from n is:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Ours is a special case where $k = 2$:

$$\binom{N}{2} = \frac{N!}{2!\,(N-2)!} = \frac{N(N-1)(N-2)!}{2!\,(N-2)!} = \frac{N(N-1)}{2}$$

Intuitively what does $\frac{N(N-1)}{2}$ mean? Well, a pair is a tuple, (a, b). For our first choice, a, we have N choices, and for our second choice, b, we have $N - 1$ choices. But for us, $(a, b) = (b, a)$ since if a has the same birthday as b, then it must be that b has the same birthday as a. Therefore to refrain from double counting, we must divide by 2.

Now given an individual a, what is the probability that another individual b has the same birthday as individual a? It should be clear that this is $\frac{1}{365}$. We can then say that that probability that b does not have the same birthday as a is $1 - \frac{1}{365} = \frac{364}{365}$.

If the probability that a and b do not have the same birthday is $\frac{364}{365}$ and the probability that a and c do not have the same birthday is $\frac{364}{365}$ then the probability that a and b *and* a and c do not have the same birthday is $\frac{364}{365} \cdot \frac{364}{365} = \left(\frac{364}{365}\right)^2$. The reader should give a justification of this fact. (The probability that there is no shared birthday among this group of 3 is $\left(\frac{364}{365}\right)^3$. A justification of this is also left for the reader.)

Given what we know now, we need to solve for the largest P that satisfies the following:

$$\left(\frac{364}{365}\right)^P \leq 0.5$$

where P is the number of pairs. Solving for P gives us the number of pairs required such that the probability that no pair of individuals share a birthday is less than 50%, which is equivalent to saying that the probability that there is a pair of individuals sharing a birthday is greater than 50%.

$$\begin{aligned}
P \cdot \ln\left(\frac{364}{365}\right) &\leq \ln(0.5) \\
P \cdot (-0.002743) &\leq -0.6931 \\
P &\leq 253
\end{aligned}$$

We therefore need at least 253 pairs. Recalling our formula to calculate the number of pairs given N individuals, we have:

$$\begin{aligned}
\frac{N(N-1)}{2} &\geq 253 \\
N(N-1) &\geq 506 \\
N^2 - N - 506 &\geq 0 \\
(N-23)(N+22) &\geq 0
\end{aligned}$$

Our value of N is therefore 23. (We disregard $N = -22$ for obvious reasons.) With 23 individuals, we can form 253 pairs of individuals, and the probability that a pair of individuals share a birthday is:

$$1 - \left(\frac{364}{365}\right)^{253} \approx 1 - 0.4995 = 0.5005$$

which is above 50% as desired.