

Problem Statement What is the probability that in a class of n students, three or more students share the same birthday? Assume there are 365 possible birthdays, each being equally likely.

Probability Space Let Ω be the set containing all possible choices and arrangements of n birthdays. We let $\mathcal{F} = 2^\Omega$ be the σ -field. For a probability measure, we choose $P(\mathcal{A}) = \frac{|\mathcal{A}|}{|\Omega|}$ for $A \subseteq \Omega$. Thus, $\{\Omega, \mathcal{F}, P\}$ is a probability space.

Pairs of students Let \mathcal{A} be the event that at least three students share the same birthday.

Let $\mathcal{B}(n, k)$ be the event that k pairs of n students uniquely sharing birthdays. First, at $k = 1$, we can use combinatorial analysis to evaluate.

If two students share the same birthday, there are $\binom{365}{1}$ ways of choosing this birthday. Next, there are $\binom{n}{2}$ ways of choosing the students who share this birthday. We use combinations and not permutations because the order of the students does not change the outcome.

For the remaining $n - 2$ people, we must choose from the remaining unpaired birthdays and distribute them to each person without replacement. Thus, there are there are $\binom{364}{n-2}$ ways of choosing the birthdays for the remaining people. Finally, because different arrangements of birthdays for the remaining unpaired people are unique outcomes, we must count the number of ways of arranging these birthdays for the remaining $n - 2$ people. This is simply $(n - 2)!$.

Putting this all together we get the number of outcomes where exactly two students share the same birthday,

$$\binom{365}{1} \binom{n}{2} \binom{364}{n-2} (n - 2)! \quad (1)$$

Thus, the probability that exactly two students share the same birthday in a group of n students is,

$$\frac{\binom{365}{1} \binom{n}{2} \binom{364}{n-2} (n - 2)!}{365^n} \quad (2)$$

If we want to instead consider the case of exactly two pairs of students, we can build on this reasoning. First, we need to choose two birthdays to be shared. Similarly, there are $\binom{365}{2}$ ways of choosing these birthdays. Next,

there are $\binom{n}{2}$ ways of choosing the first pair of students, but now, we must also consider the second pair of students. We must choose these students from the remaining $n - 2$ students. Therefore, there are $\binom{n-2}{2}$ ways of choosing the second pair of students. For the remaining $n - 4$ people, we choose the same as above; there are now $\binom{363}{n-4}$ ways of choosing from the birthdays, and $(n - 4)!$ ways of arranging these birthdays.

Putting this all together, we get that the number of outcomes where exactly four students share two distinct pairs of birthdays is,

$$\binom{365}{2} \binom{n}{2} \binom{n-2}{2} \binom{363}{n-4} (n-4)! \quad (3)$$

Thus, the probability of exactly two pairs of students sharing the same birthday is,

$$\frac{\binom{365}{2} \binom{n}{2} \binom{n-2}{2} \binom{363}{n-4} (n-4)!}{365^n} \quad (4)$$

From here, a pattern emerges. Let us assume the case of k distinct pairs of students. First we choose k birthdays, giving us $\binom{365}{k}$ ways of choosing these birthdays. Next, there are $\binom{n}{2}$ ways of choosing the first pair of students, $\binom{n-2}{2}$ ways of choosing the second pair, or, in general, $\binom{n-2m+2}{2}$ ways of choosing the m^{th} pair.

For picking the remaining birthdays, there are $365 - k$ remaining birthdays, and there are $n - 2k$ remaining unpaired people. Thus, there are $\binom{365-k}{n-2k}$ ways of choosing the remaining birthdays, and $(n - 2k)!$ ways of arranging them.

Putting this all together, we get the number of outcomes where there are k distinct pairs of birthdays among n students as,

$$\binom{365}{k} \left(\prod_{m=1}^k \binom{n-2m+2}{2} \right) \binom{365-k}{n-2k} (n-2k)! \quad (5)$$

The probability of event \mathcal{B} can be then described as,

$$P(\mathcal{B}) = \frac{\binom{365}{k} \left(\prod_{m=1}^k \binom{n-2m+2}{2} \right) \binom{365-k}{n-2k} (n-2k)!}{365^n} \quad (6)$$

Triplets of Students The original problem was finding the probability that three or more people share the same birthday. To find this probability, we try to find the complement instead. $\mathcal{A}^C = \Omega - \bigcup_k E_k$ where E_k is a pairwise disjoint subset of Ω representing the set of outcomes with exactly k pairs of students. How high is the highest possible number of pairs? If n is even, there can be $\frac{n}{2}$ possible pairs. If n is odd, it is possible to have $\frac{n-1}{2}$ pairs. In other words, it is possible to have $\lfloor \frac{n}{2} \rfloor$ pairs. Thus, $\mathcal{A}^C = \Omega - \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} E_k$

Finally, we now have the probability of at least three people sharing the same birthday.

$$P(\mathcal{A}) = 1 - P(\mathcal{A}^C) = 1 - \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{\binom{365}{k} \left(\prod_{m=1}^k \binom{n-2m+2}{2} \right) \binom{365-k}{n-2k} (n-2k)!}{365^n} \quad (7)$$

By expanding and simplifying terms, this can be written as,

$$P(\mathcal{A}) = 1 - \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{365! \prod_{m=1}^k \binom{n-2m+2}{2}}{k!(365+k-n)!365^n} \quad (8)$$