The Birthday Problem Extended

Sandra M. Pulver

It is interesting to see how the probability of at least two people in a group having the same birthday can be determined by using the classic birthday problem. The probability of success is much greater than most people would expect.

But how do we figure out more specific questions like, What are the chances of getting exactly three or four or any other number of parts of matches? This problem can be solved through an extension of the basic birthday problem to k matches.

The birthday problem is a classic example used widely in classrooms to demonstrate the principles of probability. The earliest version of this problem was developed in 1939 by Von Mises. The standard problem asks to find the probability, in a group of n individuals, that at least two of them will have the same birthday. The outcome is quite surprising because the probability of success is way beyond what most people would guess.

We compute P (at least 2 of n persons share the same birthday) as $1-P$ (no people share the same birthday).

Assuming there are 365 equally likely possibilities and n people, the number of possible birthday combinations is 365. To solve for the probability of nonmatches, we calculate the probability that each individual does not match any proceeding person.

Considering the individuals arranged in order I_1, I_2, $I_3 \ldots I_n$, I_1 can have any of the 365 birthdays, with the probability of no matches being 365/365. I_2 can have any of the 364 birthdays and not match I_1. I_3, can have any one of 363 birthdays and ... I_n can have any of $365-(n-1)$ birth dates not used.

The probability that no two people in group size n share the same birthday is

$$= \frac{365}{365} \times \frac{364}{365} \times \ldots \times \frac{(365-(n-1))}{365} = \frac{365!}{(365)^n\,(365-n)!}$$

If n is 23,the probability of no matches is 0.4927, and so, the probability of the successful outcome is $1-0.4927 = 0.5073$, which is much higher than what most people would expect.

The probability of at least one match for various n can be found in the table.

Number of People (n)	Probability of at Least One Match
5	0.0271
10	0.1169
15	0.2529
20	0.4114
23	0.5073
25	0.5687
80	0.9999

In the basic birthday problem, we solved for the probability of at least one match. This problem can be extended to a special case of Von Mises' problem:

> Determine the probability of exactly k pairs of matches in a group size n.

Let P (k) denote the probability of exactly k pairs of matches (no triple or higher matches) in a random group of size n.

Two basic principles of probability are used to develop the formula. First, the probability of an event equals the number of outcomes that meet the condition multiplied by the probability of a particular case that meet the required conditions. (This assumes that each outcome is equally likely and mutually exclusive.) For instance, the probability of 3 heads in 5 tosses of a coin that results in heads with probability p is $\binom{5}{3} p^3 (1-p)^2$ .The $p^3 (1-p)^2$ is the probability of a specific sequence and the $\binom{5}{3}$ is the number of equally likely and mutually exclusive sequences.

Second, the probability of a sequence of events is the product of each event's conditional probability.

$$P(E_1\, E_2 \ldots E_k) = P(E_1) \cdot P(E_2/E_1) \cdots P(E_k/E_1\, E_2 \ldots E_{k-1})$$

We calculate the probability of P_0 (no matches) in a group of size n.

$$P_0 = P(I_1 \neq I_{k<1}) \cdot P((I_2 \neq I_1) \cdot P(I_3 \neq I_1 I_2) \cdots P(I_n \neq I_1 I_2 \ldots, I_{n-1})$$

$$= \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-(n-1)}{365}$$

$$= \frac{365!}{(365)^n\,(365-n)!}$$

Next, we compute the probability P_1, of exactly one match. In a group of size n, there are $\binom{n}{2}$ ways of picking the individuals to have the one match and there are $\binom{365}{1}$ ways of picking the exact date to be matched.

Let P' be the probability that a particular date is matched by a specific pair in the group. For instance let P'_1 be the probability that the first two individuals were the only ones born on January 1. The probability that both I_1 and I_2 have birthdays January 1 is $(\frac{1}{365})(\frac{1}{365})$. The probability that the third person, I_3 does not match the first two individuals is $(\frac{364}{365})$, and the probability that the fourth person I_4 does not I_1 I_2 or I_3 is $(\frac{364}{365})$. Thus, the probability that the nth person does not match any of the preceding birth dates is $\frac{[365-(n-2)]}{365}$.

Therefore,

P' = P (I_1 born January 1) · P(I_2 born January 1) · P(I_3 does not match I_1 or I_2) · ... · P(I_n does not match the proceeding n − 1 individuals) =

$$= \frac{1}{365} \cdot \frac{1}{365} \cdot \frac{364}{365} \cdots \frac{[365-(n-2)]}{365}.$$

We solve for P_1 by multiplying P'_1 by the number $\binom{n}{2}$ of ways of selecting the matching pair, then by $\binom{365}{1}$ of ways of choosing the matching date.

$$P_1 = \binom{n}{2}\binom{365}{1}\frac{364!}{(365-n+1)!\,(365)^n} = \frac{\binom{n}{2}\,365!}{(365-n+1)!\,(365)^n}$$

We use these same principles to solve for P_2 of exactly two matching pairs. There are $\binom{n}{2}$ ways of choosing the first matching pair, and from the remaining n − 2 individuals, there are $\binom{n-2}{2}$ways of choosing the second matching pair. There are $\binom{365}{2}$ways of choosing the dates to be matched. Let P'_2 be the probability that I_2 and I_5 have birthdays on January 1 and I_4 and I_7 have birthdays on January 2, and no other matches exist.

P' = P(I has birth date other than January 1 or 2) · P(I was born January 1) · P(I has birth date other than January 1 or 2 and different from I) · P(I was born January 2) · P(I was born January 1) · P(I was born January 1) · P(I was not born January 1 or 2 or on the birth dates of I or I) · P(I was born on January 2) · P(I has different birthday from preceding individuals) · ... · P(I has different birth date from the proceeding individuals)

$$= \frac{363}{365} \cdot \frac{1}{365} \cdot \frac{362}{365} \cdot \frac{1}{365} \cdot \frac{1}{365} \cdot \frac{361}{365} \cdot \frac{1}{365} \cdot \frac{360}{365} \cdots \frac{(365-n+3)}{365}$$

$$= \frac{365!}{(365)^n\,(365-n+2)!}$$

To compute P_2, we multiply P'_2 by the $\binom{n}{2}\binom{n-2}{2}$ ways of selecting the individual pairs that match and then by $\binom{365}{2}$ number of ways of choosing the dates of matches.

$$P_2 = \binom{n}{2}\binom{n-2}{2}\binom{365}{2} \cdot \frac{363!}{(365-n+2)!\,(365)^n} = \frac{\binom{n}{2}\binom{n-2}{2}\,365!}{2\,(365)^n\,(365-n+2)!}$$

Finally, to calculate the probability P_k of exactly k matching pairs, we first recall that there are $\binom{n}{2}$ ways of choosing the first individuals to match, $\binom{n-2}{2}$ ways of choosing the next matching pair form the remaining (n − 2) individuals, and $\binom{n-2(k-1)}{2}$ways of picking the kth matching from the remaining n − 2(k − 1) individuals after the first (k − 1) pairs have been picked. Also, there are $\binom{365}{k}$ ways of picking the specific birthdays to be matched. Letting P'_k be the probability that I_1 and I_2 have birthdays on January 1, I_3 and I_4 have birthdays on January 2 and so on, and none of the remaining n − 2k individuals match anyone else.

P'_k = P(I_1 born January 1) · P(I_2 born January I) · P(I_3 born January 2) · P(I_4 born January 2) · ... · P(I_{2k-1} born the kth day of the year · P(I_{2k} born the kth day of the year) · P(I_{2k+1} does not match proceeding birthdays) · ... · (I_n does not match any of the proceeding birthdays)

$$= \frac{(365-k)!}{(365)^n\,(365-n+k)!}$$

There are $\binom{n}{2}\binom{n-2}{2} \ldots \binom{n-2(k-1)}{2}$ ways of picking the pairs of individuals to match birth dates and there are ways of choosing the specific dates that are matched. Thus,

$$P_k = \binom{n}{2}\binom{n-2}{2} \cdots \binom{n-2(k-1)}{2}\binom{365}{k} P'_k$$

$$= \frac{\binom{n}{2}\binom{n-2}{2} \cdots \binom{n-2(k-1)}{2}\binom{365}{k}(365-k)!}{(365-n+k)!\,(365)^n}$$

$$= \frac{n!\;365!}{(365)^n\,k!\,2^k\,(n-2k)!\,(365-n+k)}$$

= Probably of exactly k matching pairs of birthdays.

Thus, we can compute the probability of k pairs of numbers being exactly the same. For example, in a group of 100 people, the chances of having 1 match is:

$$P_1 = \frac{100!\;365!}{(365)^{100}\,1!\,2(98)!\,(365-100+1)!}$$

When this is computed, we can see that the chances of having exactly one match in a group of 100 people is extremely slim.

Bibliography

Abramson, M., and O. J. Moser. "More Birthday Surprises." *The American Mathematical Monthly* 77 (October 1970): 856–58.

Hirschhorn, M. D. "A Birthday Present for Ramanujan." *The American Mathematical Monthly* 97 (May 1990): 398–400.

Hocking, R. L., and N. C. Schwertman. "An Extension of the Birthday Problem to Exactly k Matches." *The College Mathematics Journal* 17 (January 1986): 315–21.

Jones, K. "The Birthday Problem Again?" *The Mathematics Teacher* 86 (May 1993): 373–77.

Parzen, E. *Modern Probability Theory and Its Applications*. Toronto: John Wiley, 1960.

Spencer, N. "Celebrating the Birthday Problem." *Mathematics Teacher* 70 (April 1997): 348–53.

The Circle on the Chessboard

What is the radius of the largest circle constructed on a chessboard in such a way that the perimeter of the circle lies entirely in the white squares? Where is the location of the circle's centre? The sides of the squares on the chessboard are one unit long.