ISSN 2616-7433 Vol. 3, Issue 3: 110-112, DOI: 10.25236/FSST.2021.030320

The probability principle of the birthday paradox and extended applications

Qiyang Sun

Hampton Roads Academy, 739 Academy Ln, Newport News, VA 23602, USA
Email: 22qsun@hra.org

Abstract: *There are 365 days in a year and the probability of each person's birthday is 1/365. Suppose there are n people as the observed sample and the probability of 2 people having the same birthday is 100% when n is greater than 365. When $1 \le n \le 365$, what is the probability that at least 2 people have the same birthday. The perceptual understanding of this problem is that the closer n is to 365 the greater the probability of occurrence, and the probability increases with n. Although the speed will be accelerated, it will not be outrageous. In fact, when n=3, the probability that at least 2 people have the same birthday is only 1%; The probability increases rapidly to 50.7% for n=23 and to 70.6% for n=30. The rate of probability improvement is exponentially related to the observed sample size. The problem is called the "birthday paradox" because of the obvious deviation of the computational results from the perceptual experience. It reflects the contradiction between rational calculation and perceptual understanding, not a logical contradiction.*

Keywords: *probability principle; birthday paradox; extended applications*

1. The probability calculation principle of the birthday paradox and its extension

1.1 Probability calculation principle

To make it easier to understand and calculate, we first have to start from the perspective that birthday probabilities are different. The probability of the first person's birthday is 1/365, and assuming that all people have different birthdays, the probability of the second person is 1/364 The probability of the nth person is [365-(n-1)]/365. We denote by $\bar{p}(n)$ the probability that each of the n individuals has a different birthday:

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

Conversely, we use p(n) to denote the probability that at least 2 of the n individuals have the same birthday:

$$p(n) = 1 - \bar{p}(n) = 1 - \frac{365!}{365^n(365-n)!}$$

From the equation, the probability of occurrence is 50.7% when n=23, 97% when n=50, and 99.4% when n=60. The rate of probability improvement is very impressive.

Then we calculate the probability that one of the n others has the same birthday:

$$q(n) = 1 - \left(\frac{364}{365}\right)^n$$

The probability is only about 6.1% when n = 23. If the probability of other people having the same birthday is 50% among n people, n must be at least 253.

The calculation shows that the former requires only 23 samples while the latter requires 253 samples to meet the same target probability of 50%. The birthday paradox is rooted in the fact that when we see "someone with the same birthday", we subconsciously use "someone with the same birthday" to

ISSN 2616-7433 Vol. 3, Issue 3: 110-112, DOI: 10.25236/FSST.2021.030320

speculate. The birthday paradox tells us that what is rare for an individual is common in the collective, and as the number of observed samples increases, the probability of duplication grows exponentially, and we are influenced by inertia and seriously underestimate its speed.

1.2 Birthday paradox promotion extension

We extend the probability principle of the birthday paradox to a general form. Suppose that n integers are taken from the interval [1,d], and the probability P that at least two of the data take the same value has been determined, what is the minimum value of n?

This is obtained by the equation:

$$n(p;d) \approx \sqrt{2d \ln\left(\frac{1}{1-p}\right)}$$

2. The extended application of the birthday paradox

1) Assume that the maximum value of a large venue is 1,000 people, and during the epidemic prevention and control period, the venue is divided into different small venues with separate access in order to reduce the chance of head-to-head contact between people. So that the probability of head-to-head contact between people does not exceed 20% of the 1000 people on site, how much should each small venue's reception size be controlled?

Substituting P = 20% and d = 1000 into the equation, we get

n ≈ 21.1 (persons)

That is, the staff size of each small meeting place should be controlled within 21 people in order to meet the conditions of epidemic prevention.

2) If intensive prevention and control measures are taken, masks are worn throughout the venue, effective ventilation, disinfection, etc., the size of the small venue can be expanded. If the probability of head-to-head contact can be relaxed to 90% of the 1000 people on site, what should be the size of the reception of each small venue?

Substituting P = 90% and d = 1000 into the equation, we get

n ≈ 67.8 (persons)

That is, the staff size of each small meeting place should be controlled within 67 people in order to meet the conditions of epidemic prevention.

3) Conclusion

When the probability of head-to-head human contact is relaxed to 90% for all personnel, inertia would suggest that the size should be within 900 people, even if it is more prudent to be in the hundreds. The calculation shows that the personnel should be kept within 67 people. In addition, the probability increases 3.5 times from 20% to 90%, but the staff size increases only 2.2 times, showing a non-equal proportional change. Such calculation results deviate from the results of people's senses, which also reflects the relevance of using data to speak and supporting science.

References

[1] *V.S. Vladimirov and I. V. Volovich, Teor. Mat. Fiz., 59, 3 (1984).*
[2] *I.V. Volovich, Class. Quant. Gray., 4, L83 (1987); Teor. Mat. Fiz., 71,337 (1987).*
[3] *V.S. Vladimirov, I. V. Volovich, and E. I. Zelenov, Izv. Akad. Nauk SSSR, Ser. Mat., 54, 275 (1990)o*
[4] *V.S. Vladimirov and I. V. Volovich, Commun. Math. Phys., 123, 659 (1989).*
[5] *P.G.O. Freund, M. Olson, and E. Witten, Phys. Lett. B, 199, 191 (1987).*
[6] *I. Ya. Aref'eva, B. Dragovic, and I. V. Volovich, Phys. Lett. B, 212, 283 (1988).*
[7] *A. Yu. Khrennikov, Dokl. Akad. Nauk SSSR, 314, 1380 (1990).*
[8] *A. Yu. Khrennikov, Usp. Mat. Nauk, 45, 79 (1990).*

ISSN 2616-7433 Vol. 3, Issue 3: 110-112, DOI: 10.25236/FSST.2021.030320

[9] *A. Yu. Khrennikov, J. Math. Phys., 32, 932 (1991). 10. A. Yu. Khrennikov, Teor. Mat. Fiz., 83,406 (1990).*