Parabola Volume 61 Issue 3 (2025), Article 7

Many aspects of probability theory

Ryohei Miyadera[1], **Mika Akaeda, Risana Arai, Yuga Fujiwara, Kota Inahara, Seimu Isoda, Iroha Oka, Riona Sakai, Yuito Shirataka**[2] and **Hikaru Manabe**[3]

1 Introduction

The theory of probability is a very attractive field of mathematics. It has beautiful theories and vast fields of applications. If you like mathematics and choose a topic to study, then the theory of probability is one of the best choices.

This article studies various phenomena that deal with probabilities. You can read each section independently, so you are free to skip a section and read the next section.

In Section 2, we study the well-known *Birthday Problem*. This problem captivates many people because of its surprisingly counter-intuitive answer and the simple way to calculate probabilities. It is known that the likelihood of getting two people with the same birthday will increase if the distribution of birthdays is not uniform, and here we prove this fact using mathematics understandable for high school students.

In Section 4, we study a code used in the Second World War by the Japanese Navy. Wars can change the destiny of nations, and the war to protect codes and break codes has been secretly fought behind the battles between warships, warplanes and tanks.

In Section 5, we study the distribution of blood types. Explaining the stable distribution of blood types by the Hardy-Weinberg principle is interesting because it provides a surprisingly simple and elegant mathematical framework for understanding a complex biological phenomenon. However, we could not find precise proof that the distribution of ABO blood types tends to remain stable in books and on the web. Therefore, we made a proof understandable for high school students.

In Section 6, we study the probability weighting function proposed by Daniel Kahneman and Amos Tversky. This function describes a critical feature of human psychology: we don't treat probabilities linearly. Instead, we tend to overweight small probabilities and underweight moderate to high probabilities. If you apply probability theory to society, then the probability weighting function is an essential tool. We also study the case when even corporations cannot be rational. We study the fatal mistake of the senior management that caused the nuclear meltdown during the Great East Japan Earthquake in 2011.

In Section 7, we study the probability of something spreading, such as disease and rumour.

In Section 8, we study different aspects of probability theory that deal with modern society.

[1]Ryohei Miyadera, Ph.D., is a mathematician and mathematics advisor at Keimei Gakuin in Japan (runnerskg@gmail.com).

[2]Students at Keimei Gakuin High School, Japan.

[3]Student at the University of Tsukuba, Japan.

2 The Birthday Problem

The Birthday Problem asks for the probability that there are at least two people in a group who share the same birthday. Our intuition often leads us to a minimal probability because we tend to think about the probability in relation to ourselves. We mistakenly frame the question as, "What's the chance someone has my birthday?". In reality, the problem is the probability of any two people sharing a birthday.

Suppose that there are n randomly chosen people. We calculate the probability that at least two of these people will share the same birthday.

Let the n people be numbered 1 to n. The event that all n people have different birthdays is the same as the event that person 2 does not have the same birthday as person 1, and person 3 does not have the same birthday as either person 1 or person 2, and so on. Finally, that person n does not have the same birthday as anyone from 1 to $n-1$. Then, the probability is

$$\overbrace{\frac{365}{365} \times \frac{365-1}{365} \times \cdots \times \frac{365-(n-1)}{365}}^{n} = \frac{\prod_{k=0}^{n-1}(365-k)}{365^n}, \tag{1}$$

where $\prod$ is the product operation; for precise details, see [10]. Therefore, the probability that at least two will share the same birthday is

$$f(n) = 1 - \frac{\prod_{k=0}^{n-1}(365-k)}{365^n}. \tag{2}$$

The Figure 1 is the graph of the function in (2) for $n = 1, 2, \ldots, 100$.

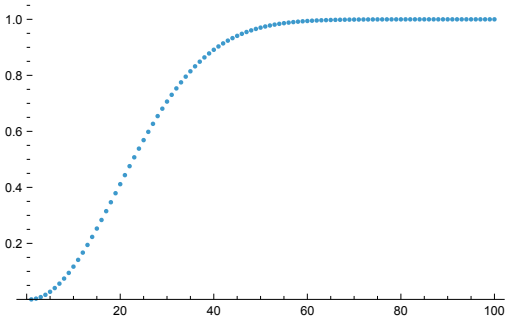

Figure 1: The Birthday Problem probabilities

Using the function in (2), we know that the probability exceeds 50% for $n = 23$ people, and most people will be surprised to learn this.

So far, we have studied the Birthday Problem under the condition that the birth rate of each day is the same in a year. However, it is known that the probability that at least two will share the same birthday is bigger if the birth rate of each day is not uniform throughout a year. We prove this fact by using mathematical knowledge that is understandable to high school students.

2.1 The probability that a random variable yields the same number

In this subsection, we study a generalized Birthday Problem. Let n be a natural number and let X be a random variable with probabilities $p_i = P(X = i)$ for $i = 1, 2, \ldots, n$. Let $U(p_1, p_2, \ldots, p_n, m)$ be the probability that X does not produce the same number in independent m trials, where $m \leq n$.

If you are a reader who does not know about random variables or independent trials, then you can think of X as a dice with n sides that might not be of the same size, such that the probability of rolling a number i is p_i for $i = 1, 2, \ldots, n$. Since there are ${}_nC_m = n(n-1)/2$ ways to choose m elements from $\{1, 2, \ldots, n\}$, there are ${}_nC_m$ subsets of $\{1, 2, \ldots, n\}$ whose size is m, and we denote these subsets by A_t for $t = 1, 2, \ldots, {}_nC_m$.

Lemma 1. *The function U can be calculated by the following equation:*

$$U(p_1, p_2, \ldots, p_n, m) = \sum_{t=1}^{{}_nC_m} \prod_{i \in A_t} p_i \,. \tag{3}$$

Proof. The probability to produce m different numbers $i_1, i_2, \ldots, i_m$ is $p_{i_1} \times p_{i_2} \times \cdots \times p_{i_m}$, and hence we have (3). □

Let i and j be natural numbers such that $1 \leq i, j \leq n$ and $i \neq j$, and let

$$\begin{aligned} T_i &= \{t : i \in A_t \text{ and } j \notin A_t\}\,; \\ T_j &= \{t : j \in A_t \text{ and } i \notin A_t\}\,; \\ T_{i,j} &= \{t : i, j \in A_t\}\,; \\ \text{and} \quad T_0 &= \{t : i, j \notin A_t\}\,. \end{aligned}$$

Also, let $q_t = p_t$ for $t \neq i, j$ and $q_i = q_j = \frac{p_i + p_j}{2}$.

Lemma 2. *The following equation holds:*

$$\sum_{t \in T_i} \prod_{u \in A_t} p_u + \sum_{t \in T_j} \prod_{u \in A_t} p_u = \sum_{t \in T_i} \prod_{u \in A_t} q_u + \sum_{t \in T_i} \prod_{i \in A_t} q_u \,. \tag{4}$$

Proof. We suppose without any loss of generality that $i = 1$ and $j = 2$ and that $A_t = \{1, 3, 4, \ldots, m+1\}$ and $A_{t'} = \{2, 3, 4, \ldots, m+1\}$. Then, $t \in T_1$, $t' \in T_2$ and

$$\begin{aligned} \prod_{u \in A_t} p_u + \prod_{u \in A'_t} p_u &= p_1 p_3 \cdots p_{m+1} + p_2 p_3 \cdots p_{m+1} \\ &= \frac{p_1 + p_2}{2} p_3 \cdots p_{m+1} + \frac{p_1 + p_2}{2} p_3 \cdots p_{m+1} \\ &= \prod_{u \in A_t} q_u + \prod_{u \in A'_t} q_u \,. \end{aligned}$$

Similarly, we prove that

$$\prod_{u \in A_t} p_u + \prod_{u \in A'_t} p_u = \prod_{u \in A_t} q_u + \prod_{u \in A'_t} q_u \tag{5}$$

for any A_t and A'_t such that $i \in A_t$, $j \in A'_t$, $j \notin A_t$, and $i \notin A'_t$. Thus, we have (4). □

Lemma 3. *The following inequality holds:*

$$\sum_{t\in T_{i,j}}\prod_{u\in A_t} p_u \le \sum_{t\in T_{i,j}}\prod_{u\in A_t} q_u\,. \tag{6}$$

Proof. We suppose without any loss of generality that $i=1$ and $j=2$, and that $A_t = \{1,2,3,4,\dots,m\}$. Then $A_t \in T_{1,2}$.

$$\prod_{u\in A_t} p_u = p_1p_2p_3\cdots p_m \le \frac{p_1+p_2}{2}\frac{p_1+p_2}{2}p_3\cdots p_m = \prod_{u\in A_t} q_u\,. \qquad \square$$

The authors made the following Lemma 4 by themselves, but they found that the result of Lemma 4 had been presented in [1].

Lemma 4. *For $i,j\in\{1,2,\dots,n\}$ such that $i\neq j$,*

$$\begin{aligned} &U(p_1,\dots,p_{i-1},p_i,p_{i+1},\dots,p_{j-1},p_j,p_{j+1},\dots,p_n,m)\\ &\quad\le U(p_1,\dots,p_{i-1},\frac{p_i+p_j}{2},p_{i+1},\dots,p_{j-1},\frac{p_i+p_j}{2},p_{j+1},\dots,p_n,m)\,. \end{aligned} \tag{7}$$

Proof. Since

$$\sum_{t\in T_0}\prod_{u\in A_t} p_u = \sum_{t\in T_0}\prod_{u\in A_t} q_u,$$

by Lemma 1, Lemma 2 and Lemma 3, we have (7). $\square$

The first author of the present article proved the following theorem in [9].

Theorem 5 ([9])**.** *Let $a_1,a_2,\dots,a_n$ be a sequence of real numbers and let $s=a_1+a_2+\cdots+a_n$. For any i such that $1\le i<n$, replace a_i and a_{i+1} by $\frac{a_i+a_{i+1}}{2}$ and $\frac{a_i+a_{i+1}}{2}$ to form a new sequence. By repeatedly using making such replacements, we eventually obtain a sequence*

$$s/n+\delta_1, s/n+\delta_2,\dots,s/n+\delta_n\,, \tag{8}$$

for real numbers $\delta_1,\delta_2,\dots,\delta_n$ such that $|\delta_1|,|\delta_2|,\dots,|\delta_n|$ can be as small as desired; i.e., each number of the sequences converges to s/n.

Theorem 6. *Let $m\le n$. The probability $1-U(p_1,p_2,\dots,p_n,m)$ of X producing the same number in m trials is minimal when $p_1=p_2=\cdots=p_n$.*

Proof. If we substitute p_i and p_{i+1} with $\frac{p_i+p_{i+1}}{2}$ and $\frac{p_i+p_{i+1}}{2}$, then by Lemma 4,

$$1-U(p_1,p_2,\dots,p_i,p_{i+1},\dots,p_n,m) \ge 1-U(p_1,p_2,\dots,\frac{p_i+p_{i+1}}{2},\frac{p_i+p_{i+1}}{2},\dots,p_n,m). \tag{9}$$

If we use this substitution repeatedly, then, by Theorem 5, these probabilities converge to the uniform probability distribution. Therefore, the function $1-U(p_1,p_2,\dots,p_n,m)$ has its minimum value when $p_1=p_2=\cdots=p_n$. $\square$

3 Real-world statistics of the Birthday Problem

The following are birthrate statistics for each day of the week in Japan in 2023 [2]:

Day	Sun	Mon	Tue	Wed	Thu	Fri	Sat
$Probability$	0.094995	0.1530370	0.1710092	0.1644052	0.1552383	0.153831	0.1074828

It is interesting to find that Sunday and Saturday are the days when the birth rates are low. In most countries where Sunday is a holiday, Sunday and Saturday are usually the days when the birth rates are low. Experts speculated that the reason why the birthrates are low on Sunday and Saturday is due to the scheduling by medical professionals.

4 Probability and breaking codes

Breaking codes used by the navy and army was an important factor in determining the outcome of the Second World War. Here, we study the case of the Japanese Navy D-code (JN-25, as called by the US). About one month before the Battle of Midway, the US could intercept about 60% of all the Japanese military communications, and the US codebreakers could read about 40% of the intercepted messages. Therefore, they could read about 24% of the whole messages. Then, by studying radio traffic patterns, identifying recurring phrases, and using IBM machines to sort data in painstaking ways, they put together fragments of information to build up the whole of the Japanese military operation. It was like assembling a jigsaw puzzle, but the situation was more difficult, because there were many missing pieces in this puzzle.

In this section, we study the mathematical aspect of cryptanalysis. JN-25 used two steps of encryption and two codebooks. The first codebook contained a 5-digit number for each word used in the war, and there was a second codebook. This second book is an "additive" book that the sender used to add to the 5-digit numbers.

Here, we need two types of number operations: addition without carrying and subtraction without borrowing.

Definition 7. *For two non-negative integers x, y that are less than 10, we define $x \oplus y$ to be the addition of y to x without carrying.*

Example 8. $3 \oplus 5 = 8$ *and* $8 \oplus 5 = 3$.

Definition 9. *For two non-negative integers x, y that are less than 10, we define $x \ominus y$ to be the subtraction of y from x without borrowing.*

Example 10. $8 \ominus 5 = 3$ *and* $3 \ominus 4 = 9$.

The two steps of encryption are Procedures 11 and 12.

Procedure 11. *Each word has a unique 5-digit number that stands for it.*

Procedure 12. *For any 5-digit number $uvwxy$, we pick up a random 5-digit number $pqrst$ from the addition book and calculate the 5-digit number $(u \oplus p)(v \oplus q)(w \oplus r)(x \oplus s)(y \oplus t)$. The location of the 5-digit random number $pqrst$ was with the message when it was sent.*

The breaking of the Japanese naval codes depended on the discovery of overlaps. The discovery of overlaps usually depended on chance but about 300 coded messages were sent every day at that time, so Americans were able to find some overlaps.

An overlap consists of messages with some common random numbers. In the overlap presented below, the first two messages share common random numbers A, B and C, and the second and third messages share common random numbers B, C and D. These random numbers A, B, C and D were chosen from the addition book.

Japanese Navy sent the 5-digit numbers that were printed in Black in the table below. The US codebreakers could read these 5-digit numbers. These 5-digit numbers were made from common random numbers of an additive book and the codes for the words by the addition without carrying.

Americans tried to find the codes for each word and common random numbers. In the table, the codes for each word and common random numbers are printed in red.

Common Random Numbers				1 A 73013	2 B 97953	3 C 03643	4 D 96867			
First Telegram Message codes	72339 *A Fleet*	44928 *20th*	37389 *leave*	37636 *22nd* 64623	97953 *arrive* 00000	26144 *plan* 23501				
Second Telegram Message codes		96337 *B Fleet*	26664 *23rd*	60598 *leave* 97585	54122 *25th* 67639	03643 *arrive* 00000	19368 *plan* 23501			
Third Telegram Message codes					70996 *C Fleet* 83403	68268 *23rd* 65625	83342 *leave* 97585	87012 *26th*	39812 *arrive*	25349 *plan*

We now study how to find the codes for each word and common random numbers, but note that this kind of code-breaking is essentially a procedure of trial and error.

The words in the overlaps, such as "arrive" and "leave", were easy to guess when the message was supposed to announce fleet departures. If they found a certain message was "23rd", then by the distance between the start point of the departure and the destination, it was not difficult to guess "25th" as the date of arrival.

Suppose that "arrive" is $x = 00000$. (We could use x in the following explanation but, by defining $x = 00000$, the explanation becomes easier to understand.) Since $B \oplus x = 97953$ and $x = 00000$, the random number is $B = 97953$. Next, we find the code y for "25th" in the second message. Since $y \oplus B = y \oplus 97953 = 54122$, $y = 54122 \ominus 97593 = 67639$. Similarly, the code for "C fleet" is $70996 \ominus 97593 = 83403$. In the second message, "arrive" is 03643. Since we suppose that the code for "arrive" is $x = 00000$, the common random number C is 03643. Then, the code for "plan" is $26144 \ominus 03643 = 23501$. In the third message, the code for "23rd" is $68268 \ominus 03643 = 65625$. Since the code for "plan" is 23501, the common random number D is $19368 \ominus 23501 = 96867$.

The above procedure of breaking Japanese Naval codes was presented in [3] by Mr. Hidemi Ito, a Kyoto University graduate who majored in theoretical physics in the Ph.D. course. He is an expert in the history of the Japanese Navy and Army codes.

4.1 The Battle of Midway and Code Breakers

The Battle of Midway, fought from June 4-7, 1942, was a pivotal naval battle in the Pacific Theater of World War II. It was a decisive victory for the United States over Japan in the Pacific. Following the attack on Pearl Harbor and a string of victories across Southeast Asia, Japan sought to eliminate the United States as a strategic Pacific power. Admiral Isoroku Yamamoto, commander-in-chief of the Japanese Combined Fleet, planned to lure the remaining U.S. aircraft carriers into a trap near Midway Atoll, a key U.S. naval base. The Japanese hoped to cripple the U.S. Pacific Fleet and force a negotiated peace. However, U.S. Navy codebreakers knew the location, timing, and even the composition of the Japanese attack force. The battle began with a Japanese air attack on Midway. However, the U.S. carriers, including the USS Enterprise, USS Hornet, and USS Yorktown, were already positioned to ambush the Japanese fleet. American dive bombers launched devastating attacks. In a matter of minutes, four Japanese aircraft carriers - Akagi, Kaga, Sōryū and Hiryū - were critically damaged and later sank. The USS Yorktown was hit and severely damaged. However, the loss was far less significant than the Japanese losses.

Behind this military success of the U.S., there was a very ingenious move by the codebreakers. Before the Battle of Midway, the U.S. noticed a high volume of traffic related to an upcoming major operation, and they also noticed that the Japanese were referring to a target with the code designation "AF". The biggest challenge was confirming the identity of "AF". While Rochefort and his team at Station HYPO suspected it was Midway, a rival intelligence unit in Washington, D.C., believed the target was in the Aleutian Islands. Admiral Chester Nimitz, the Commander in Chief of the U.S. Pacific Fleet, needed concrete proof before he could commit his limited carrier forces. To resolve the dispute, Rochefort had the naval base at Midway send out a false, unencrypted radio message, stating that the island's freshwater distillation plant had broken down and was in urgent need of water. The U.S. cryptanalysts then waited for a Japanese response. Within hours, they intercepted an encrypted Japanese message that, when deciphered, mentioned that "AF" was reporting a water shortage. This message provided the undeniable confirmation Nimitz needed.

4.2 A myth regarding codebreaking of JN-25

Some authors have written that the US Navy sank the Japanese submarine I-124, and divers were sent down and entered the submarine and removed naval codebooks, a godsend for the Navy codebreakers. In [4], Hidemi Ito, who is an expert on Japanese military codes, wrote that the capture of code books from I-124 was impossible. The US National Security Agency/Central Security Service holds the same view in [5].

5 Blood type distribution

Here, we study the distribution of blood types. It is well-known that the distribution of ABO blood types is stable in an area when there is not a large number of people coming in and going out of the area. Most people take this fact for granted but very few people know the mechanism that makes the distribution of blood types stable. Here, we use the Hardy-Weinberg principle to explain this stable distribution. We could not find a precise proof, in books or on the web, that the distribution of ABO blood types tends to remain stable. Therefore, we made a proof that is understandable for high school students.

Three factors determine the ABO blood types. Namely, factors A, B and O, and by the combination of these three factors, we have AA, $A0$, BB, BO, AB and OO types. We call AA and $A0$ *Type A*, BB and BO *Type B*, AB *Type AB*, and OO *Type O*.

Suppose that $P(A) = a$, $P(B) = b$ and $P(O) = o$. Then we have $a + b + o = 1$ and the following Lemma 13.

Lemma 13.

(i) *The probability of getting Type AA is a^2.*
(ii) *The probability of getting Type AO is $2ao$.*
(iii) *The probability of getting Type BB is b^2.*
(iv) *The probability of getting Type BO is $2bo$.*
(v) *The probability of getting Type AB is $2ab$.*
(vi) *The probability of getting Type O is 0^2.*

Proof. We get blood type AA from factor A and factor A, so we have (i). We get type AO from factor A and factor O or factor O and factor A. Hence, we have (ii). We can prove (iii), (iv), (v), (vi) similarly. □

Now, we study the distribution of blood types ABO in the next generation of people.

Theorem 14. *Suppose that no people are coming in and going out of an area, and the probability of getting married is independent of blood types. Then, the distribution of blood types remains the same in the next generation; i.e., we have the following for the next generation:*

(i) *The probability of getting Type AA is a^2.*
(ii) *The probability of getting Type AO is $2ao$.*
(iii) *The probability of getting Type BB is b^2.*
(iv) *The probability of getting Type BO is $2bo$.*
(v) *The probability of getting Type AB is $2ab$.*
(vi) *The probability of getting Type O is 0^2.*

Proof. First, we prove (ii). To calculate the probability of getting blood type AO in the next generation, we use Lemma 13. We can get AO from AA and AO, AA and BO, AA and OO, AO and AO, AO and BO, AO and AB, AO and OO, BO and AB, AB and OO. To calculate probabilities, it is better to treat the case of AA and AO as $(ii.1)$ and $(ii.2)$.

$(ii.1)$ We calculate the probability of getting AO from AA and AO. From AA and AO, we get AA, AA, OA and OA with equal probabilities and, hence, the chance of getting AO is $\frac{2}{4}$. Since the probability for blood types AA and AO is a^2 and $2ao$ respectively, the probability of getting AO from AA and AO is $\frac{2}{4} \times 2ao \times a^2 = a^3o$.

$(ii.2)$ The probability of getting AO from AO and AA is also a^3o.
$(ii.3)$ The probability of getting AO from AA and BO is $\frac{2}{4} \times a^2 \times 2bo = a^2bo$.
$(ii.4)$ The probability of getting AO from BO and AA is also a^2bo.
$(ii.5)$ The probability of getting AO from AA and OO is $\frac{4}{4} \times a^2o^2 = a^2o^2$.
$(ii.6)$ The probability of getting AO from OO and AA is also a^2o^2.
$(ii.7)$ The probability of getting AO from AO and AO is $\frac{2}{4} \times 2ao \times 2ao = 2a^2o^2$.
$(ii.8)$ The probability of getting AO from AO and BO is $\frac{1}{4} \times 2ao \times 2bo = abo^2$.
$(ii.9)$ The probability of getting AO from BO and AO is also abo^2.
$(ii.10)$ The probability of getting AO from AO and AB is $\frac{1}{4} \times 2ao \times 2ab = a^2bo$.
$(ii.11)$ The probability of getting AO from AB and AO is also a^2bo.
$(ii.10)$ The probability of getting AO from AO and OO is $\frac{2}{4} \times 2ao \times o^2 = ao^3$.
$(ii.11)$ The probability of getting AO from OO and AO is also ao^3.
$(ii.12)$ The probability of getting AO from BO and AB is $\frac{1}{4} \times 2bo \times 2ab = ab^2o$.
$(ii.13)$ The probability of getting AO from AB and BO is also ab^2o.
$(ii.14)$ The probability of getting AO from AB and OO is $\frac{2}{4} \times 2ab \times o^2 = abo^2$.
$(ii.15)$ The probability of getting AO from OO and AB is also abo^2.

By the above, we have the probability of getting type AO is

$$2a^3o+2a^2bo+2a^2o^2+2a^2o^2+2abo^2+2a^2bo+2ao^3+2ab^2o+2abo^2 = 2ao(a+b+o)^2 = 2ao\,.$$

Next, we prove (vi), by calculating the probability of getting blood type O in the next generation.

$(vi.1)$ The probability of getting OO from OO and OO is $\frac{4}{4} \times o^2 \times o^2 = o^4$.
$(vi.2)$ The probability of getting OO from OO and AO is $\frac{2}{4} \times o^2 \times 2ao = ao^3$.
$(vi.3)$ The probability of getting OO from AO and OO is also ao^3.
$(vi.4)$ The probability of getting OO from OO and BO is $\frac{2}{4} \times o^2 \times 2bo = bo^3$.
$(vi.5)$ The probability of getting OO from BO and OO is also bo^3.
$(vi.6)$ The probability of getting OO from AO and AO is $\frac{1}{4} \times 2ao \times 2ao = a^2o^2$.
$(vi.7)$ The probability of getting OO from AO and BO is $\frac{1}{4} \times 2ao \times 2bo = abo^2$.
$(vi.8)$ The probability of getting OO from BO and AO is also abo^2.
$(vi.9)$ The probability of getting OO from BO and BO is $\frac{1}{4} \times 2bo \times 2bo = b^2o^2$.

By the above, we have the probability of getting type O is

$$o^4 + 2ao^3 + 2bo^3 + a^2o^2 + 2abo^2 + b^2o^2 = o^2(a+b+o)^2 = o^2\,.$$

We can prove (iv) and (v) by a method that is similar to the one used for (ii), and we can also prove (i) and (iii) by a method that is similar to the one used for (vi). □

The result of Theorem 14 is the Hardy-Weinberg principle itself.

6 Probability and human behaviour

How do people decide their behaviour? The classical explanation uses the expected utility theory. The central premise of expected utility theory is that people choose the option with the highest expected utility, and the expected utility is

$$EU(A) = \sum_{i=1}^{n} P(A_i)U(A_i)\,,$$

where A_i is a possible outcome for $i = 1, 2, \ldots, n$, $U(A_i)$ is the utility of A_i for the person and $P(A_i)$ is the probability of A_i. Here, the concept of $U(A_i)$ is not simple. For someone, getting money is very important, although for others, it is not the case. Can the expected utility theory explain people's behaviour properly? The answer is no. In many cases, people's subjective perception of a probability deviates from the objective probability itself, and this fact explains human behaviour very well. Therefore, people often do not accept the probability $P(A_i)$ of an outcome A_i as it is. They instead distort the probability by thinking it is more than it is or thinking it is less. A. Tversky and D. Kahneman [6] proposed the probability weighting function

$$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}\,. \tag{10}$$

A probability weighting function maps objective probabilities p to decision weights $w(p)$. This function describes how individuals subjectively distort probabilities when making decisions under risk, often overweighting small probabilities and underweighting larger probabilities. Therefore, people usually use a non-linear perception of probability. Here, a non-distorted perception of probability is described by the linear function

$$L(p) = p\,. \tag{11}$$

People don't always treat probabilities linearly. They tend to be more sensitive to changes in probabilities when they are small and less sensitive when they are large. The most classic example of the probability weighting function is the purchase of lottery tickets. The odds of winning a major jackpot are astronomically low. However, millions of people still buy tickets, overestimating the likelihood of a life-changing win. Another example is people's fear of plane crashes. The objective probability of a commercial plane crash is extremely low compared to a car accident. However, due to the dramatic, catastrophic nature of a crash, it is a high-impact event. The human mind overweights this minuscule probability, making the risk feel much more significant and immediate than the statistics would suggest. People tend to forget that the consequences of car accidents can be very severe, and car accidents are prevalent.

The Birthday Problem in Section 2 also deals with the probability distorted by people's viewpoints, so we may treat it as an example of underweighting larger probabilities. In Figures 2, 3 and 4, we compare the probability weighting function (10) and the linear function (11) when $\gamma = 0.4$, $\gamma = 0.6$ and $\gamma = 0.7$.

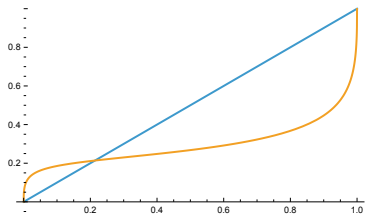

Figure 2: $\gamma = 0.4$

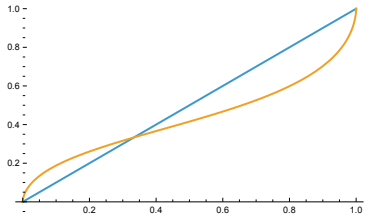

Figure 3: $\gamma = 0.6$

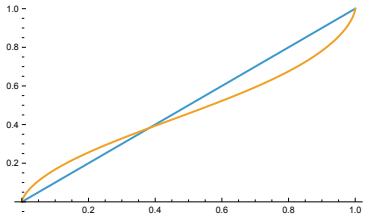

Figure 4: $\gamma = 0.7$

Then, you may ask if the probability weighting function governs the world? The answer is no. Corporations and countries, in their formal and systematic decision-making processes, largely endeavor to use the expected utility theory. Corporations are required to make decisions that maximize shareholder value. The expected utility theory, which focuses on objective probabilities and rational trade-offs, provides a formal framework for this. Corporate boards and executives are expected to justify their decisions based on a precise analysis of risks and potential returns, not on subjective biases. The specialists in corporations use tools such as stochastic optimization, risk management frameworks, and cost-benefit analysis. Therefore, corporations mainly use the expected utility theory described by the linear function in (11) in dealing with probabilities instead of using the probability weighting function. However, it is a "tendency towards" the expected utility theory, not a perfect adherence, given the human element and organizational complexities involved. Cohesive groups prioritize harmony and conformity, suppressing dissenting viewpoints, which can lead to flawed decisions. This can happen in executive committees or board meetings.

In 2008, TEPCO Sekkei, a subsidiary of TEPCO (Tokyo Electric Power Company), calculated the height of a tsunami based on a "long-term assessment" by the government's headquarters for Earthquake Research Promotion, and found that a tsunami of up to 15.7 meters, exceeding the height of the site, could hit the Fukushima Daiichi Nuclear Power Plant.

However, executives at TEPCO's head office at the time decided to postpone the construction of tsunami countermeasures based on this estimate, claiming that it would be impractical.

On March 11, 2011, the Great East Japan Earthquake hit the Fukushima Daiichi Nuclear Power Plant with a tsunami that far exceeded expectations. The tsunami caused the loss of emergency power and cooling systems, resulting in an unprecedented meltdown of the reactor core. This background was later revealed in the course of the criminal trial, and it was pointed out that there was a significant gap in perception between the engineers on site, who recognized the necessity of tsunami countermeasures within TEPCO, and the senior management, who postponed the measures.

For a reliable literature and information source of this incident, see the report of the National Diet of Japan Accident Independent Investigation Commission:

https://www.nirs.org/wp-content/uploads/fukushima/naiic_report.pdf.

Due to the reactor-core meltdown, over 100,000 people were evacuated from their homes, and many are still displaced years later. This forced displacement led to the

loss of communities, livelihoods and significant psychological and social tolls. Many older residents, in particular, were uprooted from their medical and family support systems.

The total economic cost of the nuclear accident, including compensation, decommissioning and decontamination, is estimated to be in the hundreds of billions of US dollars. Some sources place the figure at over USD 200 billion, and others suggest it could be as high as USD 300 billion.

The damage caused by the meltdown was so significant that the decision to neglect the possibility of a massive tsunami was a fatal mistake by the senior management, even if the possibility was very slim.

7 Probability of spreading

There are theories of probabilities that deal with the phenomenon of the spreading of something such as specific diseases, rumours, and so on.

7.1 Threshold

When considering the probability theory of spreading, a "threshold" is used as a criterion for an individual who receives something to pass it on to others or not. To believe a rumour, the reliability of the information must exceed a certain threshold. This reliability is determined by the credibility of the rumour itself or the relationship (trustworthiness) with the person who has passed on the rumour. For example, we can construct a model in which a rumour from a close friend is easy to believe even at a low threshold, but a rumour from a stranger is hard to believe only at a very high threshold.

We also use the threshold as a critical point for the spread of infection. The threshold is used as an essential indicator of the boundary between whether a disease spreads within a population or spontaneously converges. The *Basic Reproduction Number* R_0 is a number that indicates how many people, on average, a single infected person will transmit the disease to in a population where no one else is immune. This R_0 is one of the most representative thresholds. If $R_0 > 1$, then the disease transmission is likely to be a chain reaction, and the infection is expected to spread. If $R_0 < 1$, then the transmission is likely to converge gradually, and the disease is expected to disappear from the population. The R_0 is calculated using a probabilistic combination of parameters such as the probability of infection, duration of infection and frequency of contact. As can be seen from the above, the threshold is an essential concept for the probability of something being transmitted.

7.2 Memes

Besides rumours and diseases, other things and phenomena are transmitted. A *meme* is a unit of information or behaviour that is transmitted by imitation from one person to another in a culture or society. It is similar to the "gene" in biology, which is

the smallest unit of genetic information transmitted by an organism. Just as genes are responsible for the evolution of organisms, memes are thought to be responsible for the evolution of culture. Specifically, memes include popular words, fashion styles, popular songs and dances, certain jokes, and images and videos that are spread on the Internet. This transmission of memes is not only found in humans. The best example of a meme in animals is the chirping of the red-headed starling. Males sing only one or a few of these songs. Thus, the males were divided into dialectal groups. An experiment revealed that the singing pattern was not genetically transmitted from parent to offspring by comparing the singing patterns of fathers and sons. Individual young males adopted the chirps of other males in their territory by mimicking them, as is the case with human language. In this experiment, there were several occasions when the young male failed to imitate the old chirp and invented a new chirp. It became clear that new chirps could be produced in various ways, including changes in chirps, additions of the same chirps, and so on. New chirps and forms appeared abruptly but remained stable over the next several years. The emergence of a new chirp is described as a cultural mutation.

It is the human species that demonstrates the power of cultural evolution. Some memes are more successful than others. This is a process analogous to natural selection. If a meme is a scientific idea, then its reproduction will depend on how well the scientific community accepts it. In this case, the number of citations in scientific journals after publication can be seen as a rough measure of the survival value of the idea. In the case of the meme of popular song tunes, the degree of fecundity may be measured by the number of people whistling the song as they walk down the street. Some memes achieve remarkable short-term success through rapid proliferation but do not stay around for long. Many popular songs are examples of such phenomena.

8 Other aspects of probability theory

We now study some other aspects of probability theory.

8.1 Human error probability

Here, we introduce a practical application of probabilities in a society. Human error probability (HEP) is a concept used in human reliability analysis (HRA) to measure the likelihood of a person making a mistake while performing a specific task. This concept is fundamental in high-stakes fields like nuclear power, aviation and medicine, where human error can lead to catastrophic consequences. Here, we study human error probability in aviation because the aviation industry has been a pioneer in human reliability analysis (HRA). Human error probability (HEP) in aviation is a highly developed and critical area of study, since it is widely accepted that human error contributes to approximately $70-80\%$ of all aviation accidents. The aim of human reliability analysis is not to blame individuals who made mistakes; instead, it highlights the complex interaction between human operators, technology, the environment and the organization. Here are four major factors of accidents.

(i) Unsafe Acts of Operators: These are the direct, active failures, such as a pilot failing to follow a checklist or an air traffic controller giving an incorrect instruction.

(ii) Preconditions for Unsafe Acts: These are the conditions that predispose individuals to make errors. Examples include fatigue, stress, inadequate training, or a lack of situational awareness.

(iii) Unsafe Supervision: This level involves failures by management, such as a supervisor's failure to provide adequate guidance or a failure to correct a known problem.

(iv) Organizational Influences: These are the "latent failures" at the highest level of the organization, such as a culture that prioritizes on-time performance over safety, or a lack of resources for maintenance.

The study of human error probability begins by breaking down a complex task into a series of smaller, discrete steps. Then, they identify the conditions that can increase or decrease the likelihood of error at each step. Next, they assign error probabilities to each step by using a combination of historical data, expert judgment, and mathematical models.

The aviation industry's proactive and systemic approach to human factors has led to a significant reduction in accident rates. By focusing on understanding the causes of human error rather than simply assigning blame, the industry continues to make flying safer.

In the last part of this subsection, we would like to introduce an incident that is not widely known outside of Japan. This was not an aviation accident but this accident testified to the importance of Factor (iv) above very strongly. On April 25, 2005, at 9:19 a.m., a tragic derailment occurred on the JR Fukuchiyama Line in Amagasaki, Hyogo Prefecture, Japan. A seven-car commuter train, operated by West Japan Railway Company (JR West), derailed just before reaching Amagasaki Station. The accident happened during the morning commute and resulted in 107 deaths (106 passengers and the driver) and 562 injuries out of approximately 700 passengers on board. During the derailment, the front four cars came off the track. The first car rammed into the parking garage of an apartment building.

The Amagasaki derailment is one of the deadliest rail disasters in Japan's postwar history. The accident was primarily attributed to excessive speed. The speed limit for the curve where the derailment occurred was 70 km/h (43 mph). However, data retrieved from the train showed it was traveling at 116 km/h (72 mph). The 23-year-old driver had two infractions just before the crash: He ran a red signal, triggering the automatic train stop system (ATS). He overshot the platform at Itami Station by three carriages and had to reverse, causing a 90-second delay. By the time the train passed Tsukaguchi Station, the delay had been reduced to 60 seconds.

Investigators concluded that the driver may have been trying to make up for lost time by speeding, possibly due to fear of disciplinary consequences. The JR West culture places extreme importance on punctuality, with tight schedules and little room for delay. JR West was known for imposing a harsh disciplinary system, where drivers who committed errors were subjected to verbal abuse, manual labor, such as weeding,

forced reflection, and detailed written apologies. The driver of the ill-fated train had already been subjected to disciplinary action ten months earlier after overshooting a platform by 100 meters. Experts concluded that such disciplinary practices amounted to psychological punishment, leading to fear, stress and loss of focus among employees. He used the service brake instead of the emergency brake when he finally realized the danger, just four seconds before derailment, to avoid further infractions, since the emergency brake required official justification.

The crash highlighted gaps in urban railway safety regulations. In Japan, due to strong confidence in railway engineering, there were no strict regulations on how close residential buildings could be to train lines. As a result, the apartment building was built very close to the tracks, which amplified the human toll during the derailment. The official investigation concluded that the crash was not merely due to human error but was also the result of a rigid, punishing corporate culture and unrealistically tight train schedules.

8.2 Probability theory related to AI

Humans use mathematics to build artificial intelligence, and most of all, probability theory is an essential mathematical tool in the theory of AI. However, we need probability theory to deal with the results AI makes.

As humans began to use artificial intelligence, they began to study new types of probabilities. There are a massive number of sentences on the web and computers in the world, and some were written by humans and some were generated by AI. Therefore, it became essential to decide which was written by a human and which was not. Therefore, when a sentence is given, it is meaningful to calculate the probability that AI wrote a particular sentence. Usually, they use software that evaluates the probability that AI writes a sentence. Many teachers and professors are using this kind of software when grading student reports, and many companies also use such software to check the essays of applicants. On the internet, many AIs write articles by imitating other articles and modifying them. This kind of AI often spreads false news, so it is essential to know which articles were written by AI.

The comparison between the performance of human doctors and AI in the diagnosis of diseases is a rapidly evolving area of research. While a simple "who is better" answer is not yet definitive, several key themes and findings have emerged from recent studies.

AI systems, particularly those using machine learning and deep learning, have shown impressive accuracy in specific diagnostic tasks. For example, they excel at interpreting medical images like chest X-rays and mammograms, often achieving accuracy rates comparable to or even exceeding those of human experts. AI is also good at pattern recognition. AI is exceptionally good at recognizing subtle patterns in data that a human might miss. This is particularly useful for identifying diseases in their early stages or for detecting rare conditions. AI's ability to quickly process and analyse vast amounts of data - including medical records, imaging scans and lab results - is a significant advantage. This can lead to faster diagnoses and reduced wait times.

Human doctors possess a crucial ability that AI currently lacks: the capacity for contextual understanding. They can take into account non-verbal cues, patient histories, personal circumstances and social factors that are essential for a complete and accurate diagnosis.

While AI is good at finding patterns in large datasets, it can struggle with rare conditions or ambiguous cases that deviate from its training data. Human doctors, with their experience and clinical reasoning skills, are often better equipped to handle these complexities.

Studies have shown that AI struggles with the nuanced and natural process of interviewing a patient to gather crucial diagnostic information. Human doctors are far superior at conducting these conversations and asking the right follow-up questions.

The doctor-patient relationship is built on trust and empathy, qualities that are currently impossible for AI to replicate. Human doctors provide emotional support and personalized care that is a vital part of the healing process.

Acknowledgements

Prof. Kit Yates's fascinating book [11] has inspired the authors. They would like to thank him very much.

References

[1] G.C. Berresford, The uniformity assumption in the Birthday Problem, *Mathematics Magazine* **53** 1980, 286–288.

[2] E-Stat, *Number of births, place of birth, place of birth, month of birth, date of birth, by birth date, and time of birth*, Japan in Statistics, https://www.e-stat.go.jp/dbview?sid=0003411915, last accessed on 2025-09-16.

[3] H. Ito, *The Defeat of Japanese Naval Codes – How the "D" Cipher Was Broken. Vol. 1*, Shiho-Shuppan, 2020.

[4] H. Ito, *The Defeat of Japanese Naval Codes – How the "D" Cipher Was Broken. Vol. 2*, Shiho-Shuppan, 2020.

[5] JN-25, Historical Events, National Security Agency/Central Security Service, https://www.nsa.gov/History/Cryptologic-History/Historical-Events/Article-View/Article/2740680/jn-25/, last accessed on 2025-09-14.

[6] D. Kahneman and A. Tversky, Prospect Theory: An analysis of decision under risk, *Econometrica* **47** (1979), 263–291.

[7] U. Malmendier and G. Tate, CEO overconfidence and corporate investment, *The Journal of Finance* **60** (2005), 2661–2700.

[8] T.S. Nunnikhoven, A Birthday Problem solution for nonuniform birth frequencies, *The American Statistician* **46** (1992), 270–274.

[9] R. Tanaka, J. Miyamoto, Y. Maruo, K. Nakayama and R. Miyadera, An elementary proof that the regular polygon is the largest among polygons that are inscribed in a circle, *Parabola* **59(1)**, 2023.

[10] Wikipedia, *Product (mathematics)*, https://simple.wikipedia.org/wiki/Product_(mathematics), last accessed on 2025-09-14.

[11] K. Yates, *The Math of Life and Death: 7 Mathematical Principles That Shape Our Lives (Using Math in Everyday Life)*, Scribner, 2021.