Mathematical Statistics

Kenichi ISHIKAWA

http://ishiken.free.fr/lecture.html

Oct. 18 Combination and Probability
Oct. 25 Random variables and probability distributions
Nov. 1 Representative probability distributions
Nov. 8 (First half) Random walk and the gambler's ruin problem
(Latter half) Brownian motion and diffusion
Nov. 22 Noise theory

Reference book
Junkichi SATSUMA "Probability /Statistics" -- Beginning of mathematics of science and technology course 7, 2001, Iwanami Shoten

Only exercises are provided on Nov. 15.

Reference books

[1] Satsuma, J. (2001). *"Probability /Statistics"* -- *Beginning course of mathematics of science and technology 7*, Iwanami Shoten

[2] Kolmogorov, A.N., Žurbenko, I. G., Prokhorov, A.V. (2003). *"Beginning of Kolmogorov's Probability Theory"*. (Trans. Murayama , T. & Baba, Y.): Morikita Shuppan

[3] KITAHARA, K. (1997). *"Nonequilibrium Statistical Mechanics"* – *Iwanami Fundamental Physics Series 8*, Iwanami Shoten

Mathematical Statistics

Kenichi ISHIKAWA

Oct. 18 Probability

- Permutations and combinations
- One-dimensional random walk (Brownian motion)
- Definition of probability
- Properties of probability
- Conditional probability

1-1 Permutations and combinations

→ Reference book [1]

- Prepare 26 cards on which each letter of the alphabet is written. Put all the cards into a bag. Then randomly take them out one by one. At this time, what is the probability that any three successive cards show B, I, S in a row?
- Prepare 11 cards on which the letters A, F, I, I, M, N, N, O, O, R, and T are written. Then put all the cards into a bag, and randomly take them out one by one. At this time, what is the probability that the cards form the word INFORMATION in order?

⇒ Permutations

- Permutations
  - Permutation is defined as an ordering of a certain number of objects of a given set.

The number of permutations that r different objects randomly chosen from among n different objects are strung in a row is:

$$ {}_nP_r = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!} $$

In particular, the number of permutations that all n different objects are strung in a row is:

$$ {}_nP_n = n(n-1)(n-2)\cdots 2\cdot 1 = n! $$

[Example] ${}_nP_r = {}_{n-1}P_r + r\,{}_{n-1}P_{r-1}$ ⟹ Recurrence formula relating the case of n objects to the case of n-1 objects

- Repeated Permutations
  - The number of permutations (Repeated Permutations), which allow choosing r objects from n different objects and stringing them in a row by repeatedly using the same objects, is n^r.
  - [Example] When constructing triple-digit natural numbers by using four numeric figures of 1, 2, 3, and 4, according to the product rule, the total number thereof is $4\times 4\times 4 = 4^3 = 64$.

- Permutation in the case when the same objects exist
  - [Example] Prepare 11 cards on which such the letters A, F, I, I, M, N, N, O, O, R, and T are written. Then put all the cards into a bag, and randomly take them out one by one. At this time, what is the number of combinations of words (letters) that can be formed by stringing such cards? (It is not necessary to consider whether such words thus formed have meanings being included in a dictionary.)

Given that n objects are divided into c sets, and that the objects belonging to the same set cannot be distinguished, but those belonging to different sets can be distinguished, the number of permutation stringing all those n objects in a row is:

$$\frac{n!}{n_1!n_2!\cdots n_c!} \qquad (n_1 + n_2 + \cdots + n_c = n)$$

If each set has one single object, then the permutation becomes the ordinary type.

- Combinations

Combination is defined as a set of objects chosen from given multiple objects without ordering.

The number of combinations of r objects randomly chosen from n different objects is:

$$ {}_nC_r = \frac{{}_nP_r}{r!} = \frac{n(n-1)\cdots(n-r+1)}{r(r-1)\cdots 2\cdot 1} = \frac{n!}{r!(n-r)!} $$

It is also often represented as $\binom{n}{r}$.

[Example] The number of combinations to form a set of three balls chosen from five balls having of different colors is:

（５×４×３）÷（３×２×１）＝１０

[Example] $ {}_nC_r = {}_{n-1}C_r + {}_{n-1}C_{r-1} $ ⟹ Recurrence formula relating the case of n objects to the case of n-1 objects

$$ {}_2C_r = \cdots $$

$$ {}_3C_r = \cdots $$

$$ {}_4C_r = \cdots $$

- Binomial theorem

Formula to represent the expansion of $(a+b)^n$

$$(a+b)^2 = a^2 + 2ab + b^2$$
$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^n = \underbrace{(a+b)(a+b)\mathcal{L}\,(a+b)}_{n}$$

The coefficient of $a^{n-r}b^r$ is the combination ${}_nC_r$ of choosing r bs from n factors $(a+b)$.

When n is a positive integer,

$$(a+b)^n = \sum_{r=0}^{n} {}_nC_r a^{n-r} b^r$$

Newton's binomial formula

Accordingly, ${}_nC_r = \binom{n}{r}$ is also referred to as the binomial coefficient.

[Example] Coefficient of x^3y^5 in the expansion of $(5x\text{-}3y)^8$ is:

- Repeated Combinations

[Example] If you buy three bottles of wine at a liquor shop selling two types of wines (red and white), there are four combinations of possible purchase combination that is three reds, two reds and one white, one red and two whites, and three whites.

⇒ Segmentation by inserting a “partition” between the sets of red and white
[Example] ○○｜○＝two reds, one white

The number of combinations of repeatedly taking r objects out of n different objects:

$$ {}_nH_r = {}_{n+r-1}C_r = \frac{n(n+1)\cdots(n+r-1)}{r!} $$

- Combinations of dividing objects into several subsets

[Example] Dividing seven students into two subsets comprising three and four each

⇒ Just simply align all, and divide the seven into the front three and the rear four. Since the order of each three and four is no object, the number of combinations is 7!/ (3!×4!)=35.

The number of combinations of dividing n different objects into c sets of n_1, n_2, …, n is:

$$ \frac{n!}{n_1!n_2!\cdots n_c!} \qquad (n_1+n_2+\cdots+n_c=n) $$

Same as “Permutation in the case when the same objects exist”

- Polynomial theorem

Combinations of objects divided into several sets

$$(a_1 + a_2 + \cdots + a_m)^n = \sum \frac{n!}{n_1! n_2! \cdots n_m!} a_1^{n_1} a_2^{n_2} \cdots a_m^{n_m}$$

However, the sum is taken over all $n_1, n_2, \cdots, n_m$ meeting $n_1 + n_2 + \cdots + n_m = n$, where $n_1 \geq 0, n_2 \geq 0, \cdots, n_m \geq 0$.

[Example] The coefficient of $x y^2 z^3$ in the expansion of $(x - 5y + 3z)^6$ is:

1-2 One-dimensional random walk (Brownian motion)

→ Books of reference [2]

Applications of probability theory are not limited to mere subjects related to dice and cards！

- **Kinetic theory of gases**
- **Diffusion phenomenon**
- **Noises**
- **Stock quote/foreign exchange**
- …

Kinetic theory of gases, Diffusion phenomenon, Noises: Nonequilibrium statistical mechanics

Stock quote/foreign exchange: Mathematical Finance

All of the above: Stochastic Processes

Chaotic and unregulated movement of individual particles ⟺ Clear and simple regularity on the whole

Probability theory（Theory of stochastic processes ）

▪Brownian Motion

Robert Brown, English botanist, 1827

Brown observed that grains of pollen suspended in water performed an erratic, uninterrupted movement.

The figure is omitted due to copyright.

（Source: J. Perrin's experimental validation; Reference book [2], p.7）

Life force of grains of pollen?

Subsequently it was proved that every grain, small enough, has such general properties.

[Example] Diffusion of an ink drop in the aquarium water

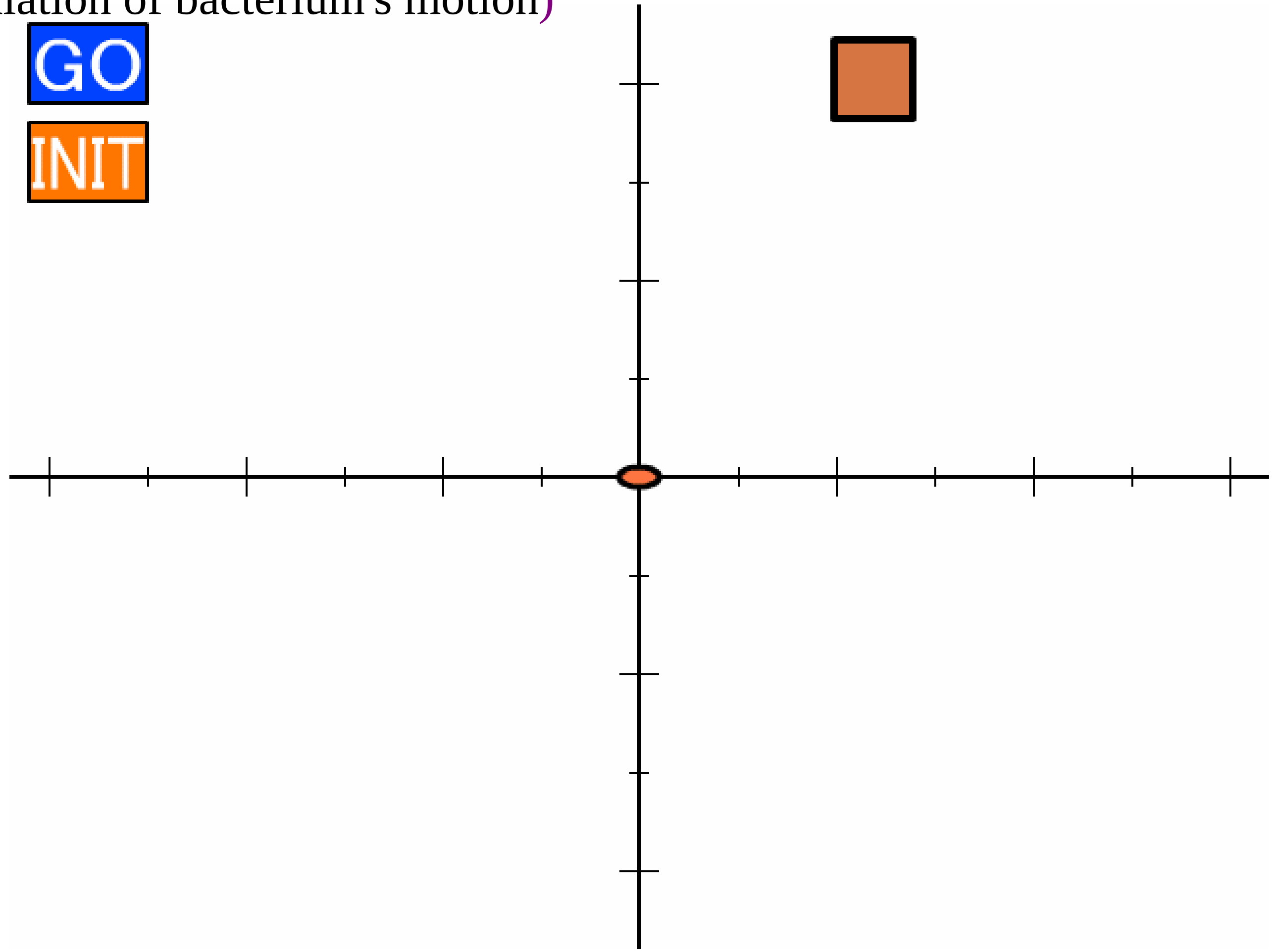
$$r = a\sqrt{t}$$

Radius of an ink drop is not proportional to time.

Radius of an ink drop

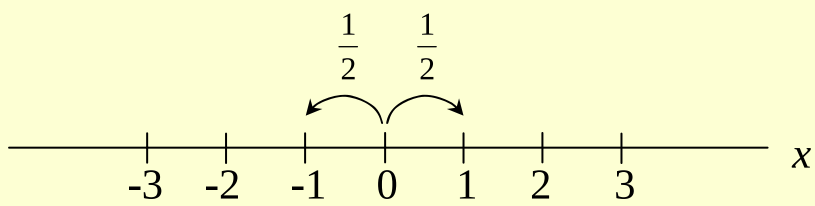
・Two-dimensional random walk (Brownian motion)

(simulation of bacterium's motion)

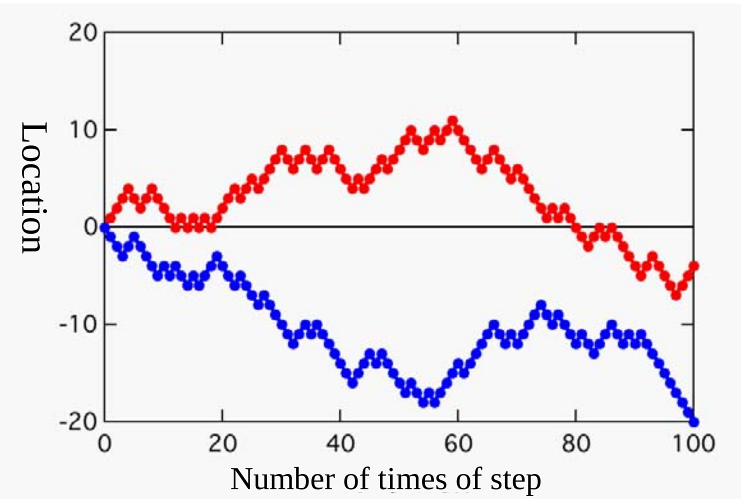


Created by Yuuki Kishida, BIS forth-year grade

▪One-dimensional random walk (Brownian motion)



- A particle leaves m=0 at time t=0
- One-time jump allows one-space move to the right or left.
- Irrespective of location, the probability that the particle moves to the right or left is one-half.



Location of a particle after N steps: $m(N)$

$$m(0) = 0$$

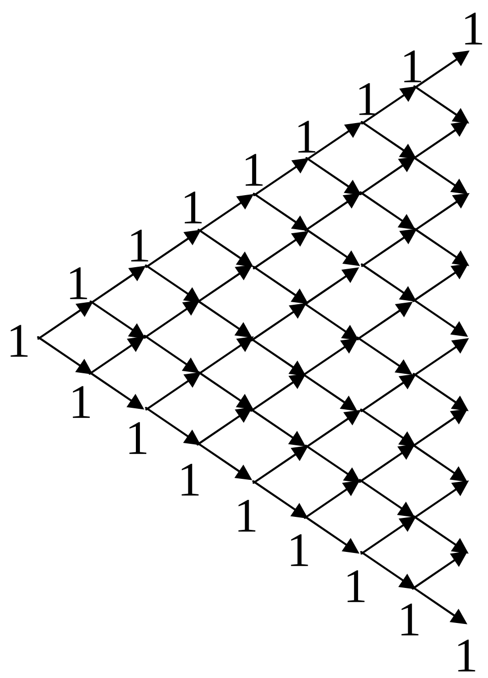
$$m(1) = -1,\ 1$$

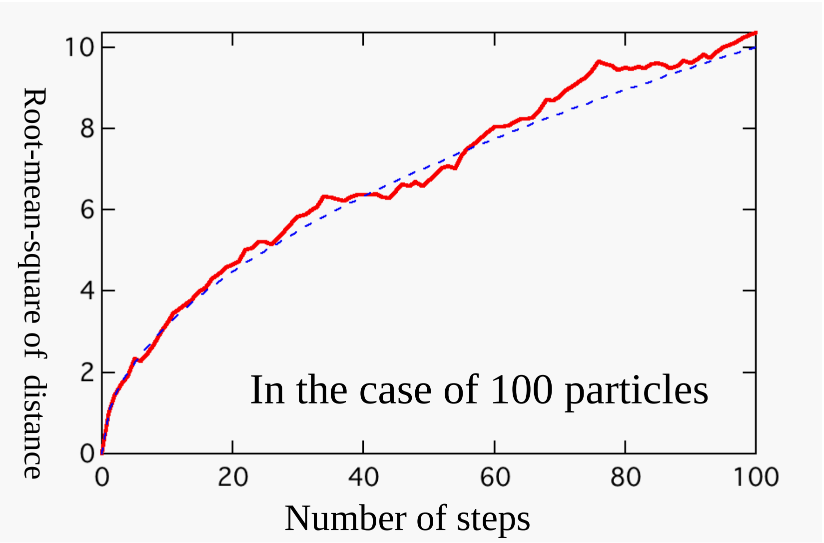
$$m(2) = -2,\ 0,\ 2$$

$$m(3) = -3,\ -1,\ 1,\ 3$$

$$m(N) = -N, -N+2, -N+4, \cdots, N-4, N-2, N = N-2k \ (k = 0, 1, \cdots, N)$$

・Calculation of the number of trajectories in straight lines





According to the binomial theorem, $\sum_{k=0}^{n} {}_nC_k = 2^n$ $\sum_{k=0}^{n} k\,{}_nC_k = n2^{n-1}$ $\sum_{k=0}^{n} k^2\,{}_nC_k = n(n+1)2^{n-2}$

Find the root-mean-square of distance by using these formulae.

1-3 Definitions of probability

→ Reference book [1]

- Objects of probability/statistics
  - Repeatable action under the conditions considered identical : (eg., casting dice, tossing coin, voltage measurement of electrical noises, etc.)
  - Agglomerations of a large number of individual objects having the same quality : (eg., large number of balls with the same size, gas molecules within a vessel, etc.)
- Trials: Operations such as casting dice to read the number thereon, measuring the kinetic energy of gas molecules, etc.
- Events: Outcomes obtained by performing trials
  - (Eg., getting three spots, or odd-number of spots on a die)
- Sample space: Set of all possible outcomes
  - Elementary event: An event not further divided (eg, getting three spots, etc.)
  - Combined event: An event including more than two elementary events (eg, getting an odd-number of spots on a die)

Through these events, it is uncertain to know which event could be obtained by performing each single trial.

↓

Increasing the number of times of trials, there may exist some regularity.

⇒ Theoretical studies of probabilities

1-3 Definition of probability

- Mathematical probability
  - Developed by P. L. Laplace

For a trial, suppose that the size of a sample space is n, and that every elementary event is likely to occur at the same rate. When an event E is taken in the sample space, and the number of possible occurrence of E is r, the probability $P(E)$ of E is defined as:

$$P(E) = \frac{r}{n}$$

（Reference book [1], p. 19）

[Example] What is the probability of getting six heads and four tails, when tossing ten 100-yen coins? Here, however, suppose that both heads and tails are likely to occur at the same rate.

[Example] Toss a 100-yen coin 10 times repeatedly, and score ○ for a head and × for a tail in a notebook. Which probability of getting (A) or (B) score shown below is higher, and how much higher than the other?

(A) ○○○○○○○○○○

(B) ○××○○×○○×○

1-3 Definition of probability

- Experimental probabilities（statistical probability）
  - Batting averages of baseball, probability of weather forecasts proving correct, appearance of numbers of spots of imperfect dice, etc.

Suppose that an event E occurred r times after performing trials n times. If r/n approximates a certain value p when increasing n, then let the probability $P(E)$ of E be:

$$P(E) = p = \lim_{n\to\infty} \frac{r}{n}$$

（Reference book [1], p. 21）

[Example] Suppose a batter (in a baseball game) with a batting average of .333 made an out in the first and second at-bat. What is the probability that the batter gets a hit in his third at-bat? (Reference book [1], p. 21）

[Example] Casting perfect dice several times repeatedly, six spots came up five times in a row. Which probability is higher in the next casting of the dice: that six spots come up, or that one spot comes up?

[Example] There are 40 students in a department. What is the probability that the birthday of one student coincides with that of another student?

1-4 Properties of probability →Reference book [1]

Here, collective concepts are used.

- Given that a sample space is S, *this* S is a set, and the event E is a subset of S. $E \subset S$
- **Product events** of A and B : Events in which both events A and B occur simultaneously $A \cap B$
- **Sum of sets** of A and B : Events in which at least one of the event A and B occurs $A \cup B$
- **Complementary events** of E: Events in which E does not occur within S $\overline{E}$
- **Empty events** ϕ : Events that never occur
- **Exclusive**: When the event A *and* B do not occur simultaneously, it is said that A and B are "exclusive" of one another.
  - [Example] Given that A is even numbers of spots and B is five-spots in casting dice, A and B are exclusive of one another. $A \cap B = \phi$
  - Every elementary event is exclusive of one another.

- Probability axioms → Reference book [1], p. 23
  - The probability $P(E)$ of each event E in a sample space S satisfies the three conditions shown below.
  - When the real number $P(E)$ exists that satisfies the three conditions shown below for each event E in a sample space S, $P(E)$ is considered the probability of occurrence of the event E.

(1) $0 \le P(E) \le 1$

(2) $P(S) = 1,\ P(\phi) = 0$

(3) When $E_1, E_2, E_3, \cdots$ are the events exclusive of one another:

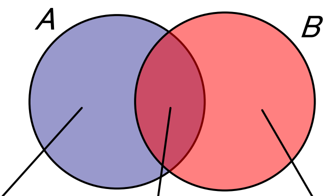
$$P(E_1 \cup E_2 \cup E_3 \cup \cdots) = P(E_1) + P(E_2) + P(E_3) + \cdots$$

Even an infinite number of quantity of events will do.

- Some formulae derived from the probability axioms → Reference book [1], p. 24

Addition Formula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



$E_1 = A \cap \overline{B}$

$E_2 = \overline{A} \cap B$

$E_3 = A \cap B$

[Example] What is the probability that a card taken from 52 well-shuffled cards excluding the joker is a spade (A) or a court card (B)?

$$P(A) = \frac{13}{52},\ P(B) = \frac{12}{52},\ P(A \cap B) = \frac{3}{52} \Rightarrow P(A \cup B) = \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{11}{26}$$

$$P(\overline{E}) = 1 - P(E)$$

⇐ $E \cup \overline{E} = S,\ E \cap \overline{E} = \phi$

[Example] When the probability of rain is 70%, the probability of no rain is: 1−0.7=0.3 (ie., 30%).

東京大学工学部 システム創成学科 C コース Program for Biological & Information Systems

1-5 Conditional probability

- Conditional probability → Reference book [1] p.27

Given that there are two events (*A* and *B*), the event that *B* will occur on condition that A has occurred is represented by $B|A$. In addition, the probability thereof $P(B|A)$ is referred to as the conditional probability of *B* under the condition *A*, and is defined by the following formula:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

No need to think about the case in which it was not a spade card.

[Example] Taking one of cards

A : Event in which it is a spade card

B : Event in which it is a court card

$B|A$: Event in which it is a court card at the same time as being a spade card.

$$P(B|A) = \frac{3}{13}\left(= \frac{3/52}{13/52} = \frac{P(A \cap B)}{P(A)} \right)$$

- Multiplication Theorem → Reference book [1], p. 27

Conditional probability of B under the condition A: $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$

Conditional probability of A under the condition B: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Multiplication Theorem $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$

[Example] Lot drawing → Reference book [1] p.27

When there are ten slips, let the event that the first drawer draws a winning slip be A, and the event that the second drawer does so be B.

When there is one winning slip

$$P(A) = \frac{1}{10} \qquad P(B) = P(A \cap B) + P(\bar{A} \cap B) = 0 + P(\bar{A})P(B|\bar{A}) = \frac{9}{10} \cdot \frac{3}{9} = \frac{3}{10}$$

When there are two winning slips

$$P(A) = \frac{2}{10} = \frac{1}{5} \qquad P(B) = P(A \cap B) + P(\bar{A} \cap B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$

$$= \frac{2}{10} \cdot \frac{1}{9} + \frac{8}{10} \cdot \frac{2}{9} = \frac{1}{5}$$

• Bayes's Theorem (Thomas Bayes)

[Example] A drug check system indicates positive reactions of steroids users at a 98% rate of probability, and also indicates positive reaction of non-steroid users at a 10% rate of probability. In a succor[soccer] club, 20 percent of players use steroids, and one player tested positive for a drug through the check. What is the probability that this player uses steroids?

Probability of causes / Posterior probability ⇨ Genetic research; Spam mails detection (Bayesian filter)

Viscerally consider it.

Use the multiplication theorem to consider the above probabilities.

Bayes's Theorem When an outcome E derives from n causes of $A_1, A_2, \cdots, An$ that are exclusive of each other, satisfying every case, the probability $P(A_i|E)$ that A_i is a cause among them is:

$$P(A_i \mid E) = \frac{P(A_i)P(E \mid A_i)}{P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2) + \cdots + P(A_n)P(E \mid A_n)}$$

→ Reference book [1], p. 30

- Bayes's Theorem → Reference book [1] p.30

Bayes's Theorem When an outcome E derives from n causes of $A1, A2, \cdot\cdot, An$ that are exclusive of each other, satisfying every case, the probability $P(Ai|E)$ that Ai is a cause among them is:

$$P(A_i \mid E) = \frac{P(A_i)P(E \mid A_i)}{P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2) + \cdots + P(A_n)P(E \mid A_n)}$$

[Example] Suppose that there are three types of machines, A, B, and C, and that A makes up 10% of total production of these machines, B makes up 30% thereof, and C 60%. In addition, suppose that the defect production rate(E) of A is 3%, that of B is 2%, and C 1%. Under the above assumption, when you sample a machine, and find it defective, the probability that such defective machine belongs to A can be calculated through the following formula:

$$P(A \mid E) = \frac{P(A)P(E \mid A)}{P(A)P(E \mid A) + P(B)P(E \mid B) + P(C)P(E \mid C)}$$

$$= \frac{0.1 \times 0.03}{0.1 \times 0.03 + 0.3 \times 0.02 + 0.6 \times 0.01} = 20\%$$

→ Reference book[1], p.31

- Statistical Independence

A and B are statistically independent. $P(A|B)=P(A)$

⇕ Multiplication theorem

$P(A\cap B)=P(A)P(B)$

Suppose that there are n events from $A_1, A_2, \cdots$ to A_n.

If $(2\le k\le n)$ $P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$ is held for any number of the events $Ai1, Ai2, \cdots, Aik$ taken from $A1, A2, \cdots$ to An, then, the events $A_1, A_2, \cdots, A_n$ are statistically independent of each other.

[Example] Suppose that there are three types of machines, A, B, and C, and that A makes up 10% of total production of these machines, B makes up 30% thereof, and C 60%. In addition, suppose that the defect production rate (E) of A is 8%, that of B is 2%, and C 1%. When you sample a machine and find it defective, the probability that such defective machine belongs to B can be calculated through the following formula:

$$P(B|E)=\frac{0.3\times 0.02}{0.1\times 0.08+0.3\times 0.02+0.6\times 0.01}=30\%$$

In addition, $P(B)=30\%$

⇒ *B and E* are statistically independent. Caution ! *A* and *C* are not statistically independent of *E*.