概 率 论

Probability Theory

概率 Probability——可能性

概率　Probability———可能性

- There will probably be at least one section of that course offered next year.

概率　Probability——可能性

- There will probably be at least one section of that course offered next year.
- 人人有机会得100分，

概率 Probability——可能性

- There will probably be at least one section of that course offered next year.
- 人人有机会得100分,
- It is likely that the Dow-Jones average will increase by the end of the year,

概率　Probability——可能性

- There will probably be at least one section of that course offered next year.
- 人人有机会得100分,
- It is likely that the Dow-Jones average will increase by the end of the year,
- It is expected that at least 20000 concert tickets will be sold,

概率 Probability——可能性

- There will probably be at least one section of that course offered next year.
- 人人有机会得100分,
- It is likely that the Dow-Jones average will increase by the end of the year,
- It is expected that at least 20000 concert tickets will be sold,
- ....

Chance, expectation, ..., these words we are using frequently are related to probability.

Chance, expectation, ..., these words we are using frequently are related to probability.

Probability theory originated from calculations of probabilities in games of chance.

Where does the chance come from? If every thing is deterministic, we have no chance.

Where does the chance come from? If every thing is deterministic, we have no chance.

There are normally two classes of phenomena in nature and human society.

- Deterministic phenomenon
- Random phenomena

Deterministic phenomenon(确定性现象), is bound to take place under a certain condition.

Deterministic phenomenon(确定性现象), is bound to take place under a certain condition. For example

1. The area in a circle is $S = \pi r^2$;

Deterministic phenomenon(确定性现象), is bound to take place under a certain condition. For example

1. The area in a circle is $S = \pi r^2$;
2. The distance a free-fall object travels within t minutes is $gt^2/2$;

Deterministic phenomenon(确定性现象), is bound to take place under a certain condition. For example

1. The area in a circle is $S = \pi r^2$;
2. The distance a free-fall object travels within t minutes is $gt^2/2$;
3. Water boils at 100 centigrade under standard atmosphere pressure.

Deterministic phenomenon(确定性现象), is bound to take place under a certain condition. For example

1. The area in a circle is $S = \pi r^2$;

2. The distance a free-fall object travels within t minutes is $gt^2/2$;

3. Water boils at 100 centigrade under standard atmosphere pressure.

sure event (Ω) and impossible events ($\emptyset$)

数学家喜欢用精确的数学公式描述自然自然界和人类社会中的规律：

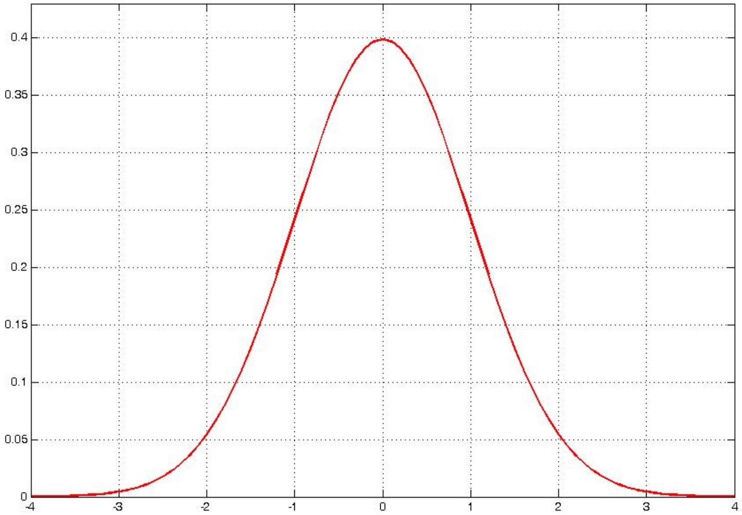
$$y = f(x)$$

数学家喜欢用精确的数学公式描述自然自然界和人类社会中的规律:

$$y = f(x)$$

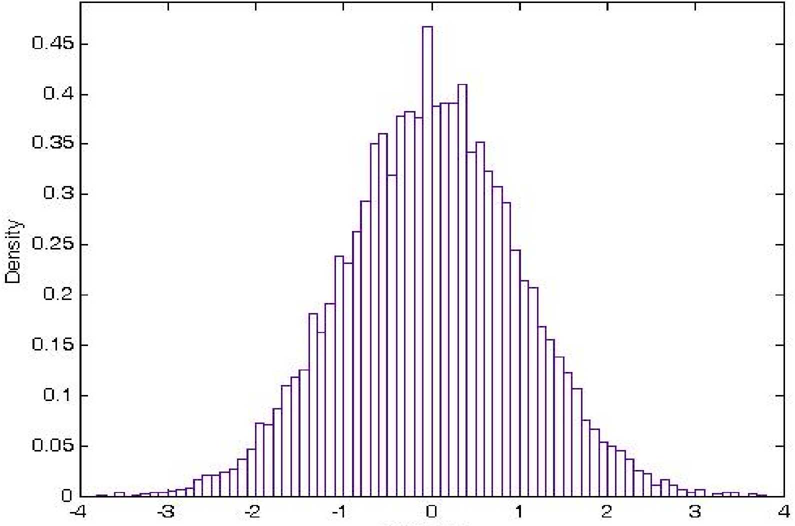
概率统计学家倾向于用带有随机误差项的数学公式描述自然界和人类社会中的规律：

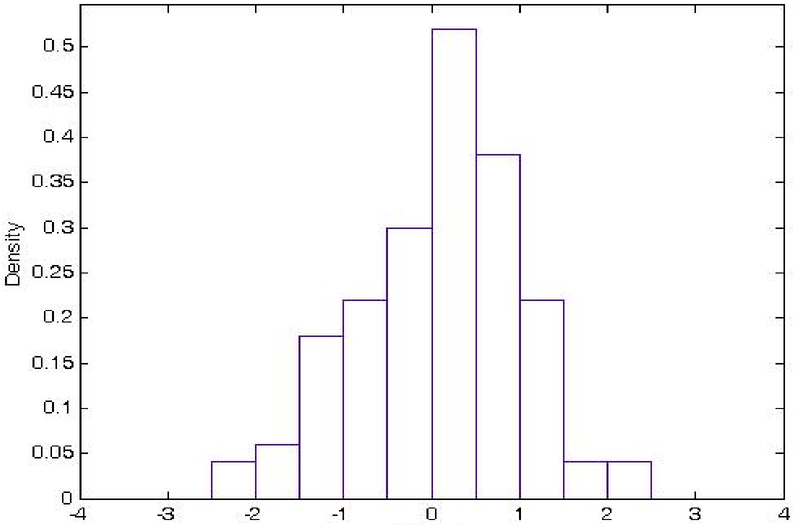
$$y = f(x) + \epsilon$$

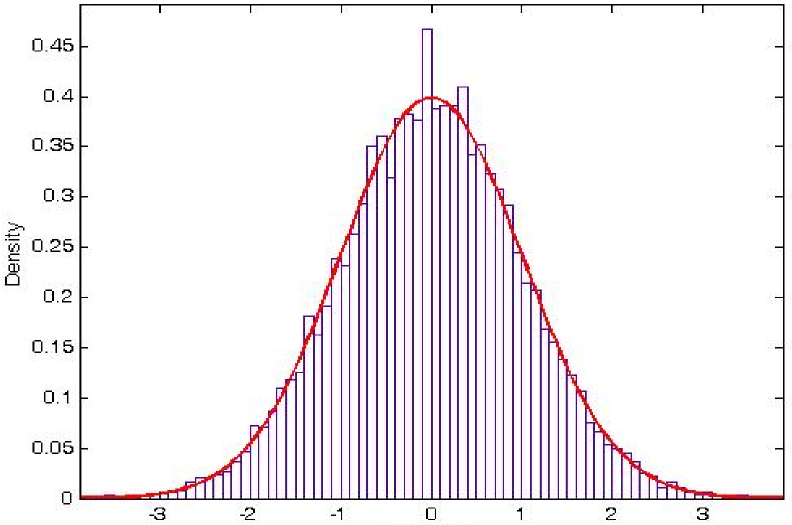

0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-4
-3
-2
-1
0
1
2
3
4

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$-\infty < x < \infty$$


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Density
-4
-3
-2
-1
0
1
2
3
4


Density
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-4
-3
-2
-1
0
1
2
3
4


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Density
-3
-2
-1
0
1
2
3

Random phenomenon(随机现象), in which the events are not sure to take place.

Random phenomenon(随机现象), in which the events are not sure to take place.

For example

1. It will rain on next January 1 in Hangzhou;

Random phenomenon(随机现象), in which the events are not sure to take place.

For example

1. It will rain on next January 1 in Hangzhou;
2. 850 out of 1000 seeds one plants will germinate;

Random phenomenon(随机现象), in which the events are not sure to take place.

For example

1. It will rain on next January 1 in Hangzhou;
2. 850 out of 1000 seeds one plants will germinate;
3. The distance between the target and the projectile one fires is less than 15 meters.

Random phenomenon(随机现象), in which the events are not sure to take place.

For example

1. It will rain on next January 1 in Hangzhou;
2. 850 out of 1000 seeds one plants will germinate;
3. The distance between the target and the projectile one fires is less than 15 meters.

random event(随机事件)

Properties of random phenomenon (random events)

Properties of random phenomenon (random events)

(1) uncertainty(不确定性)

In a random phenomenon, the outcome is not predicted with certainty in each individual experiment.

If we repeatedly observe a random phenomenon under a certain condition, the outcomes may be different from places to places.

Remark 1： But all possible outcomes are known.

Remark 1： But all possible outcomes are known.

Remark 2. A random phenomenon can be repeated under the same condition.

(2) statistical regularity(统计规律性)

The outcomes of a great deal of repeated experiments under a certain condition turn out to have some regularity (statistical regularity)

—The probability of each outcome to occur is deterministic.

Probability theory is a branch of mathematics that focuses on the study of quantity regularity of various random phenomena.

Probability theory is a branch of mathematics that focuses on the study of quantity regularity of various random phenomena.

Universal existence of random phenomena determines the importance of this subject.

Various applications: Probability theory can be used in almost all disciplines and professions, such as

- statistics;
- operations research;
- biology;
- economics;
- psychology
- ....

2019职业排行(CareerCast网站)

	Jobs	Overall Rating	Median Salary	Projected Growth
1	Data Scientist	97	$114,520	19%
2	Statistician	110	$84,760	33%
3	University Professor	112	$76,000	15%
7	Information Security Analyst	126	$95,510	28%
8	Mathematician	127	$84,760	33%
9	Operations Research Analyst	128	$81,390	27%
10	Actuary	141	$101,560	22%

2018职业排行(CareerCast网站)

	Jobs	Rated Score	Income	Growth Outlook
1	Genetic Counselor	100	$74,120	29%
2	Mathematician	102	$81,950	33%
3	University Professor	105	$75,430	15%
5	Statistician	111	$84,060	33%
7	Data Scientist	121	$111,840	19%
8	Information Security Analyst	126	$92,600	28%
9	Operations Research Analyst	127	$79,200	27%
10	Actuary	135	$100,610	22%

2017职业排行(CareerCast网站)

	Jobs	Rated Score	Income	Growth Outlook
1	Statistician	93.00	$80,110	34%
3	Operations Research Analyst	102.00	$79,200	30%
4	Information Security Analyst	104.00	$90,120	18%
5	Data Scientist	105.00	$111,267	15.75%
6	University Professor	110.00	$72,416	15.24%
7	Mathematician	128.00	$111,298	22.31%
8	Software Engineer	129.00	$100,690	17%

2016职业排行(CareerCast网站)

	Jobs	Rated Score	Income	Growth Outlook
1	Data Scientist	91.00	$128,240	16%
2	Statistician	96.00	$79,990	34%
3	Information Security Analyst	94.00	$88,890	18%
6	Mathematician	126.00	$103,720	21%
7	Software Engineer	131.00	$97,990	17%
8	Computer Systems Analyst	133.00	$82,710	21%
10	Actuary	138.00	$96,700	18%

2015职业排行(CareerCast网站)

	Jobs	Rated Score	Income	Growth Outlook
1	Actuary	80.00	$94,209	25.09%
2	Audiologist	88.00	$71,133	33.33%
3	Mathematician	92.00	$102,182	25.91%
4	Statistician	96.00	$79,191	25.91%
5	Biomedical Engineer	117.00	$89,165	25.65%
6	Data Scientist	121.00	$124,149	14.97%
8	Software Engineer	129.00	$93,113	21.13%
10	Computer Systems Analyst	135.00	$81,150	23.50 %

The definition of probability

The statistical definition of probability

Frequency:

If event A occurs n_A times in N repeated experiments under a certain conditions, then frequency of A occurring in N experiments is defined as:

$$F_N(A) = \frac{n_A}{N}.$$

When N is large enough, the frequency turns out to have a kind of stability,

i.e., the values of $F_N(A)$ show fluctuations which become progressively weaker as N increases, until ultimately $F_N(A)$ stabilizes to a constant.

Example: In tossing a fair coin, let

$A = \{\text{head comes up}\}$.

Example: In tossing a fair coin, let

$A = \{\text{head comes up}\}$.

Experimenter	Number of tosses	Times of head	Frequency
Buffon	4040	2048	0.5069
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005

Example: In tossing a fair coin, let

$A = \{\text{head comes up}\}$.

Experimenter	Number of tosses	Times of head	Frequency
Buffon	4040	2048	0.5069
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005

Frequency stabilizes to–$P(A) = \frac{1}{2}$

The statistical definition of probability:

The constant to which the frequency of the event A stabilizes calls the probability of the occurrence of event A (the probability of A).

The statistical definition of probability:

The constant to which the frequency of the event A stabilizes calls the probability of the occurrence of event A (the probability of A).

Example:The use frequency of each English letter:

E—0.105; ....; J, Q, Z—0.001

Recall the frequency of event A occurring in N experiments is defined as:

$$F_N(A) = \frac{n_A}{N}.$$

Recall the frequency of event A occurring in N experiments is defined as:

$$F_N(A) = \frac{n_A}{N}.$$

1. $F_N(A) \geq 0$;

Recall the frequency of event A occurring in N experiments is defined as:

$$F_N(A) = \frac{n_A}{N}.$$

1. $F_N(A) \geq 0$;
2. $F_N(\Omega) = 1$, where Ω is a sure event;

Recall the frequency of event A occurring in N experiments is defined as:

$$F_N(A) = \frac{n_A}{N}.$$

1. $F_N(A) \geq 0$;
2. $F_N(\Omega) = 1$, where Ω is a sure event;
3. Suppose that A and B will never come up simultaneously and that $A + B$ stands for the event that A, or B, or both come up, then $F_N(A + B) = F_N(A) + F_N(B)$.

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;
2. Normalization(规范性): $P(\Omega) = 1$, where Ω is a sure event;

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;
2. Normalization(规范性): $P(\Omega) = 1$, where Ω is a sure event;
3. Additivity(可加性): Suppose that A and B will never come up simultaneously and that $A + B$ stands for the event that A, or B, or both come up, then $P(A + B) = P(A) + P(B)$.

1.2 Classical probability models(古典概率模型)

1.2 Classical probability models(古典概率模型)

Random experiment(随机试验), Sample points(样本点) and sample spaces(样本空间)

1.2 Classical probability models(古典概率模型)

Random experiment(随机试验), Sample points(样本点) and sample spaces(样本空间)

- We roll a dice and the possible outcomes are 1, 2, 3, 4, 5, 6 corresponding to the side that turns up.

1.2 Classical probability models(古典概率模型)

Random experiment(随机试验), Sample points(样本点) and sample spaces(样本空间)

- We roll a dice and the possible outcomes are 1, 2, 3, 4, 5, 6 corresponding to the side that turns up.
- We toss a coin with possible outcomes H (heads) and T (tails).

Each elementary outcome of a random experiment is termed a sample point, usually denoted by ω.

Each elementary outcome of a random experiment is termed a sample point, usually denoted by ω. The set of all sample points is termed sample space, usually denoted by Ω.

Each elementary outcome of a random experiment is termed a sample point, usually denoted by ω.

The set of all sample points is termed sample space, usually denoted by Ω.

Let $\omega_i = \{i \text{ comes up}\}$ in the proceeding example of dice throwing, then $\Omega = \{\omega_1, \omega_2, \cdots, \omega_6\}$.

Example 1. A bag contains 10 balls, 3 of which are red, 3 white and 4 black. If a ball is drawn at random, then the sample space may be taken as

$$\Omega_1 = \{\text{a red ball, a white ball, a black ball}\}.$$

Example 1. A bag contains 10 balls, 3 of which are red, 3 white and 4 black. If a ball is drawn at random, then the sample space may be taken as

$$\Omega_1 = \{\text{a red ball, a white ball, a black ball}\}.$$

If we label these 10 balls with, red–$1, 2, 3$, white–$4, 5, 6$, black– $7, 8, 9, 10$, and take a ball randomly, then the sample space may be taken as

Example 1. A bag contains 10 balls, 3 of which are red, 3 white and 4 black. If a ball is drawn at random, then the sample space may be taken as

$$\Omega_1 = \{\text{a red ball, a white ball, a black ball}\}.$$

If we label these 10 balls with, red–$1, 2, 3$, white–$4, 5, 6$, black– $7, 8, 9, 10$, and take a ball randomly, then the sample space may be taken as

$$\Omega_2 = \{\omega_1, \cdots, \omega_{10}\}, \quad \omega_i = \{\text{ the } i\text{-th ball }\}$$

Example 2. In the above Example 1, if two balls are taken at a time, then each sample point is expressed as (i, j), where i and j are the numbers attached to balls,

Example 2. In the above Example 1, if two balls are taken at a time, then each sample point is expressed as (i, j), where i and j are the numbers attached to balls, and so the sample space is equal to

$$\Omega = \{(1,2),(1,3),\cdots,(1,10),$$
$$(2,3),\cdots,(2,10),\cdots,(9,10)\},$$

which consists of $\binom{10}{2} = 45$ sample points in sum.

Example 4. The distance between the target and the projectile is a non-negative real number. In this case the sample space may be taken as

$$\Omega = [0, \alpha],$$

a 1-dimensional continuous interval, where α is a positive real constant.

Classical probability models: possesses the following two fundamental features.

1. The sample space is finite, i.e., $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$, where $\omega_i, i = 1, 2, \cdots, n$ are elementary events.
2. Each elementary event comes up with equal possibility, i.e., their probabilities are identical.

Definition 1 If a random experiment possesses n elementary events of equal possibility and event A contains m of these elementary events, then the probability $P(A)$ that event A comes up is defined as

$$P(A) = \frac{m}{n} = \frac{\#A}{\#\Omega}.$$

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;
2. Normalization(规范性): $P(\Omega) = 1$, where Ω is a sure event;

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;
2. Normalization(规范性): $P(\Omega) = 1$, where Ω is a sure event;
3. Additivity(可加性): Suppose that A and B will never come up simultaneously (i.e., $A \cap B = \emptyset$) and that $A + B$ stands for the event that A, or B, or both come up(i.e., $A \cup B$), then $P(A + B) = P(A) + P(B)$.

In particular, if event A and B will never come up simultaneously (i.e., $A \cap B = \emptyset$), and, at least one of A and B will come up (i.e., $A \cup B = \Omega$), then

$$P(A) = 1 - P(B).$$

Example 6 有 n 个球, N 个格子 $(n \leq N)$, 球与格子都是可以区分的. 每个球落在各格子内的概率相同(设格子足够大, 可以容纳任意多个球). 将这 n 个球随机地放入 N 个格子, 求:

(1) 指定的 n 格各有一球的概率;

(2) 有 n 格各有一球的概率.

Solution.

把球编号为$1 \sim n$, n个球的每一种放法是一个样本点, 每一种放法是等可能的, 这属于古典概率模型,

把球编号为$1 \sim n$, n个球的每一种放法是一个样本点, 每一种放法是等可能的, 这属于古典概率模型, 样本点总数应该是

把球编号为$1 \sim n$, n个球的每一种放法是一个样本点, 每一种放法是等可能的, 这属于古典概率模型, 样本点总数应该是N个中取n个球的重复排列数N^n.

把球编号为$1 \sim n$, n个球的每一种放法是一个样本点, 每一种放法是等可能的, 这属于古典概率模型, 样本点总数应该是N个中取n个球的重复排列数N^n.

(1) 记$A = \{$指定的n格各有一球$\}$,

把球编号为$1\sim n$, n个球的每一种放法是一个样本点, 每一种放法是等可能的, 这属于古典概率模型, 样本点总数应该是N个中取n个球的重复排列数N^n.

(1) 记$A=\{$指定的n格各有一球$\}$, 它包含的样本点数是指定的n格中n个球的全排列数$n!$,

把球编号为$1\sim n$, n个球的每一种放法是一个样本点, 每一种放法是等可能的, 这属于古典概率模型, 样本点总数应该是N个中取n个球的重复排列数N^n.

(1) 记$A=\{$指定的n格各有一球$\}$, 它包含的样本点数是指定的n格中n个球的全排列数$n!$,

$$P(A)=n!/N^n.$$

(2)记$B = \{$有n格各有一球$\}$,

(2)记$B=\{$有n格各有一球$\}$, 它包含的样本点数是N格中选取n格的排列数P_N^n,

(2)记$B=\{$有n格各有一球$\}$, 它包含的样本点数是N格中选取n格的排列数P_N^n,

$$
\begin{aligned}
P(B) =& \frac{P_N^n}{N^n} = \frac{N!}{N^n(N-n)!} \\
=& \frac{N(N-1)\cdots(N-n+1)}{N^n} \\
=& (1-\frac{1}{N})(1-\frac{2}{N})\cdots(1-\frac{n-1}{N}).
\end{aligned}
$$

注意到

$$\log(1-x) = -x + O(x^2); \; x \to 0.$$

注意到

$$\log(1-x) = -x + O(x^2); \ x \to 0.$$

我们有

$$\log\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)$$

$$= \sum_{k=1}^{n-1} \log\left(1-\frac{k}{N}\right) = -\sum_{k=1}^{n-1}\frac{k}{N} + O\left(\sum_{k=1}^{n-1}\frac{k^2}{N^2}\right)$$

$$= -\frac{n(n-1)}{2N} + O\left(\frac{n^3}{N^2}\right) \ \text{if} \ \frac{n}{N} \to 0.$$

所以

$$P(B) \approx \exp\left\{ -\frac{n(n-1)}{2N} \right\}.$$

例 (生日问题) 求n个人中至少有两个人同生日的概率.

例 (生日问题) 求n个人中至少有两个人同生日的概率.

解.认为每个人的生日等可能地出现在365天中的任意一天, 则样本空间Ω的元素个数为$\sharp\Omega = 365^n$.

例 (生日问题) 求n个人中至少有两个人同生日的概率.

解.认为每个人的生日等可能地出现在365天中的任意一天, 则样本空间Ω的元素个数为$\sharp\Omega = 365^n$. 用$\overline{A}$表示n个人的生日各不相同, 则作为Ω的子集$\sharp\overline{A} = P_{365}^n$.

例 (生日问题) 求n个人中至少有两个人同生日的概率.

解. 认为每个人的生日等可能地出现在365天中的任意一天, 则样本空间Ω的元素个数为$\sharp\Omega = 365^n$. 用$\overline{A}$表示n个人的生日各不相同, 则作为Ω的子集$\sharp\overline{A} = P_{365}^n$. 要求的概率为

$$
\begin{aligned}
p_n =& P(A) = 1 - P(\overline{A}) \\
=& 1 - \frac{P_{365}^{n}}{365^n} = 1 - \frac{365!}{(365-n)!365^n}.
\end{aligned}
$$

$$\begin{aligned} p_n =& P(A) = 1 - P(\overline{A}) \\ =& 1 - \frac{P_{365}^n}{365^n} = 1 - \frac{365!}{(365-n)!365^n}. \end{aligned}$$

计算这样的概率, 要计算阶乘数$n!$. 随n的增大, $n!$增长非常快, 例如

$$10! = 3,628,000, \quad 15! = 1,307,674,368,700,$$

而100!包含158位数字.

在实际计算中, 常常用Stirling 公式进行近似计算:

$$
\begin{aligned}
n! =& \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left\{\frac{\theta_n}{12n}\right\}, \quad 0 < \theta_n < 1 \\
\approx& \sqrt{2\pi n}\left(\frac{n}{e}\right)^n .
\end{aligned}
$$

另一方面,

$$
\begin{aligned}
p_n =& 1 - \frac{P_{365}^n}{365^n} \\
=& 1 - \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\cdots\left(1 - \frac{n-1}{365}\right) \\
\approx& 1 - \exp\left\{ - \frac{n(n-1)}{2 \times 365} \right\} \widehat{=} \widetilde{p}_n.
\end{aligned}
$$

可以计算出如下结果:

n	20	30	40	50	60	70	80
p_n	0.411	0.706	0.891	0.970	0.994	0.9992	0.9999
$\widetilde{p}_n$	0.406	0.696	0.882	0.965	0.992	0.9987	0.9998

n	22	23
p_n	0.4757	0.5073
$\widetilde{p}_n$	0.4689	0.5000

$$p_n - \widetilde{p}_n \leq 0.01.$$

Example 7. A bag contains a white balls and b black balls. These balls are drawn one by one randomly and without replacement. Find the probability that the k-th ball drawn is a white one.

Solution 1. 把球编号, 按摸球的次序把球排成一列, 直到$(a+b)$个球都摸完, 每种排法为一个样本点, 样本点总数为

Solution 1. 把球编号, 按摸球的次序把球排成一列, 直到$(a+b)$个球都摸完, 每种排法为一个样本点, 样本点总数为$(a+b)$个球的排列数$(a+b)!$.

Solution 1. 把球编号, 按摸球的次序把球排成一列, 直到$(a+b)$个球都摸完, 每种排法为一个样本点, 样本点总数为$(a+b)$个球的排列数$(a+b)!$. 所考察的事件相当于在第k个位置放白球, 共有a种放法,

Solution 1. 把球编号, 按摸球的次序把球排成一列, 直到$(a+b)$个球都摸完, 每种排法为一个样本点, 样本点总数为$(a+b)$个球的排列数$(a+b)!$.
所考察的事件相当于在第k个位置放白球, 共有a种放法, 每种放法又对应其它$(a+b-1)$个球的$(a+b-1)!$放法,

Solution 1. 把球编号, 按摸球的次序把球排成一列, 直到$(a+b)$个球都摸完, 每种排法为一个样本点, 样本点总数为$(a+b)$个球的排列数$(a+b)!$.
所考察的事件相当于在第k个位置放白球, 共有a种放法, 每种放法又对应其它$(a+b-1)$个球的$(a+b-1)!$放法,故该事件包含的样本点数为$a(a+b-1)!$, 所求的概率为

$$P=\frac{a(a+b-1)!}{(a+b)!}=\frac{a}{a+b}.$$

Solution 2. 各球不编号, 即所有白球都看成相同, 这时相当于在$(a+b)$个位置中取a个位置放白球, 共有

Solution 2. 各球不编号, 即所有白球都看成相同, 这时相当于在$(a+b)$个位置中取a个位置放白球, 共有$\binom{a+b}{a}$中取法,

Solution 2. 各球不编号, 即所有白球都看成相同, 这时相当于在$(a+b)$个位置中取a个位置放白球, 共有$\binom{a+b}{a}$中取法, 所考察的事件为在第k个位置放白球, 其它各位置放$a-1$个白球, 共$\binom{a+b-1}{a-1}$种放法,

Solution 2. 各球不编号, 即所有白球都看成相同, 这时相当于在$(a+b)$个位置中取a个位置放白球, 共有$\binom{a+b}{a}$中取法, 所考察的事件为在第k个位置放白球, 其它各位置放$a-1$个白球, 共$\binom{a+b-1}{a-1}$种放法,

故所求的概率为

$$P=\frac{\binom{a+b-1}{a-1}}{\binom{a+b}{a}}=\frac{a}{a+b}.$$

Example 8 There are a defective products and b nondefective products and they are indistinguishable. If n $(n \leq a)$ products are sampled from them, find the probability that the n products sampled contain k defective ones.

Solution. 在$(a+b)$件产品中取n件有$\binom{a+b}{n}$种取法,

Solution. 在$(a+b)$件产品中取n件有$\binom{a+b}{n}$种取法, 而在a件次品中取k件, b件正品中取$n-k$件共有$\binom{a}{k}\binom{b}{n-k}$种取法. 故所求的概率为

$$P = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}.$$

Example 9. One has two boxes of matches, each having n matches, in his pocket. Each time he wants to use match, he will randomly take out a box and draw one match from it. When he finds the box he takes out is empty, find the probability that the other box has just m matches.

Solution. It is obvious that the event { When one box he takes out is empty, the other box has just m matches } is equal to the event { At the $(2n+1-m)$-th draw, he finds one box is empty }.

Solution. It is obvious that the event { When one box he takes out is empty, the other box has just m matches } is equal to the event { At the $(2n+1-m)$-th draw, he finds one box is empty }. And also, it is equal to $A+B$, where

$$A = \{ \text{ at the } (2n+1-m)\text{-th draw,}$$

$$\text{he finds box A is empty}\},$$

$$B = \{ \text{ at the } (2n+1-m)\text{-th draw,}$$

$$\text{he finds box B is empty}\}.$$

We consider A first.

We consider A first. It is obvious that

$A = \{$ in the first$(2n+1-m)$ draws,

box A is drawn at the $(2n+1-m)$-th draw;

and in other $(2n-m)$ draws , box A is drawn n times,

box B is drawn $n-m$ times$\}$.

We consider A first. It is obvious that

$A = \{$ in the first$(2n+1-m)$ draws,

box A is drawn at the $(2n+1-m)$-th draw;

and in other $(2n-m)$ draws , box A is drawn n times,

box B is drawn $n-m$ times$\}$.

We have totally 2^{2n+1-m} ways to take the first $(2n+1-m)$ draws in which $\binom{2n-m}{n}$ ways satisfy the condition in event A.

Similarly, in the totally 2^{2n+1-m} ways to take the first $(2n+1-m)$ draws, $\binom{2n-m}{n}$ ways satisfy the condition in event B. Therefore, the desired probability is

$$\frac{2\binom{2n-m}{n}}{2^{2n+1-m}} = \frac{\binom{2n-m}{n}}{2^{2n-m}}.$$

古典概率模型的推广

在古典概率模型中, 样本空间$\Omega = \{\omega_1, \cdots, \omega_n\}$是有限的且每个样本点出现是等可能的.

古典概率模型的推广

在古典概率模型中, 样本空间$\Omega = \{\omega_1, \cdots, \omega_n\}$是有限的且每个样本点出现是等可能的.

一般地, 如果样本空间$\Omega = \{\omega_1, \omega_2, \cdots\}$含有可列个元素, 样本点$\omega_i$出现的可能性为$p(\omega_i)$, 其中$p(\omega_i) \geq 0$, $\sum_{i=1}^{\infty} p(\omega_i) = 1$. 这时事件$A$的概率为

$$P(A) = \sum_{i:\omega_i \in A} p(\omega_i).$$

古典概率模型的推广

在古典概率模型中, 样本空间$\Omega = \{\omega_1, \cdots, \omega_n\}$是有限的且每个样本点出现是等可能的.

一般地, 如果样本空间$\Omega = \{\omega_1, \omega_2, \cdots\}$含有可列个元素, 样本点$\omega_i$出现的可能性为$p(\omega_i)$, 其中$p(\omega_i) \geq 0$, $\sum_{i=1}^{\infty} p(\omega_i) = 1$. 这时事件$A$的概率为

$$P(A) = \sum_{i:\omega_i \in A} p(\omega_i).$$

对这样的概率模型,容易验证有如下性质:

1. 非负性: $P(A) \geq 0$;

1. 非负性: $P(A) \geq 0$;
2. 规范性: $P(\Omega) = 1$;

1. 非负性: $P(A) \geq 0$;
2. 规范性: $P(\Omega) = 1$;
3. 可加性:若事件A和B不同时发生,
   则$P(A+B) = P(A) + P(B)$;

1. 非负性: $P(A) \geq 0$;
2. 规范性: $P(\Omega) = 1$;
3. 可加性:若事件A和B不同时发生,
   则$P(A+B) = P(A) + P(B)$;
4. 可列可加性:若事件A_i, $i = 1, 2, \cdots$中任何两个都不会同时发生(即,$A_i \cap A_j = \emptyset$, $i \neq j$), 用$\sum_{i=1}^{\infty} A_i$表示它们中至少有一个发生(即$\bigcup_{i=1}^{\infty} A_i$), 则
$$P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Geometrical probability models(几何概率)

Sample space Ω—a region in $\mathbf{R}^n$.

"equal possibility":

Geometrical probability models(几何概率)

Sample space Ω—a region in $\mathbf{R}^n$.

"equal possibility": measure of A = measure of B

$\Longrightarrow P(A) = P(B)$.

Definition Event $A_g = \{$ a sample point falls into region $g \subset \Omega\}$. The probability of A_g is defined as

$$P(A_g) = \frac{\text{Measure of } g}{\text{Measure of } \Omega}.$$

This is called the geometric probability.

Example 11 (The arrangement problem). Two people make an appointment to meet at a park between 7 o'clock and 8 o'clock and the person who first arrives at the park will keep waiting for another for 20 minutes. Find the probability that they can meet.

Solution......

Take 7 o'clock as the beginning time and assume that one people arrives at x and the other arrives at y. The sample ponit is (x, y) and

Take 7 o'clock as the beginning time and assume that one people arrives at x and the other arrives at y. The sample ponit is (x, y) and the sample space is

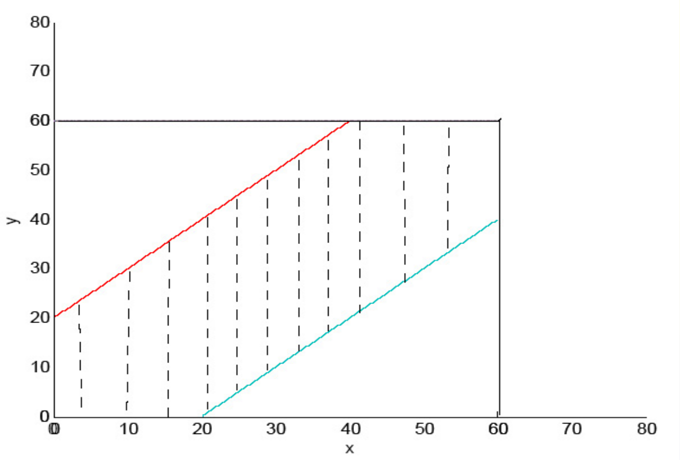
$$\Omega = \{(x, y) \big| 0 \le x \le 60, 0 \le y \le 60\}.$$

Take 7 o'clock as the beginning time and assume that one people arrives at x and the other arrives at y. The sample ponit is (x, y) and the sample space is

$$\Omega = \{(x, y) \big| 0 \leq x \leq 60, 0 \leq y \leq 60\}.$$

The two people meet each other if and only if $|x - y| \leq 20$. Therefore the sample points such that event $A = \{$they meet each other$\}$ happens constitute the area

$$g = \{(x, y) \big| |x - y| \leq 20, 0 \leq x, y \leq 60\}.$$


80
70
60
50
40
30
20
10
0
y
0
10
20
30
40
50
60
70
80
x

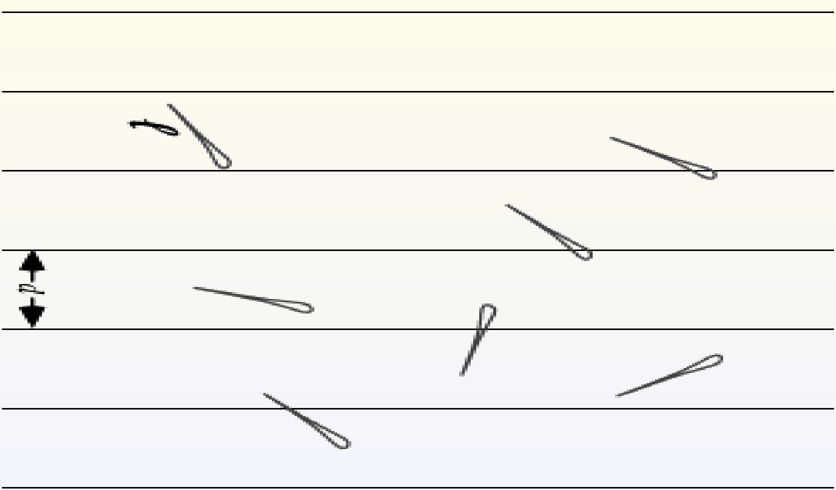
So we have

$$P(A) = \frac{\text{the area of } g}{\text{the area of } \Omega}$$

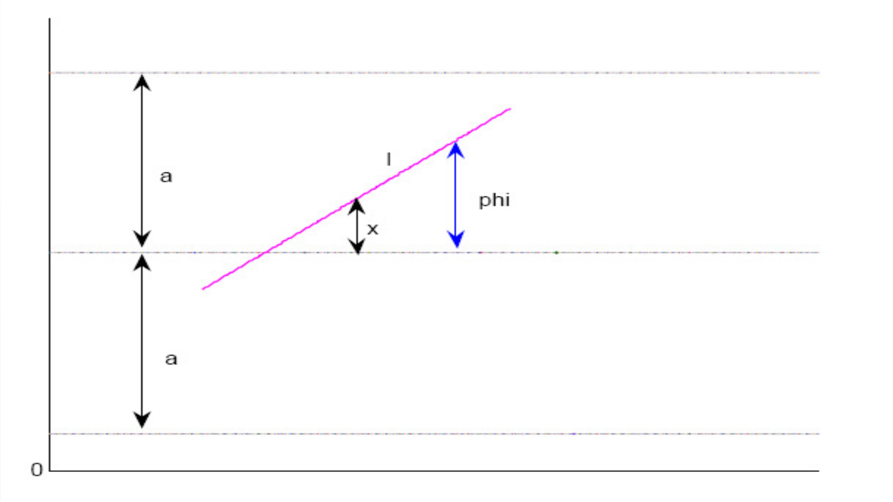
So we have

$$P(A) = \frac{\text{the area of } g}{\text{the area of } \Omega} = \frac{60^2 - (60-20)^2}{60^2} = \frac{5}{9}.$$

Example 12. (The problem of Buffon's needles) If a needle of length l is dropped at random on the middle of a horizontal surface ruled with parallel lines a distance $a > l$ apart, what is the probability that the needle will cross one of the lines?



Solution.

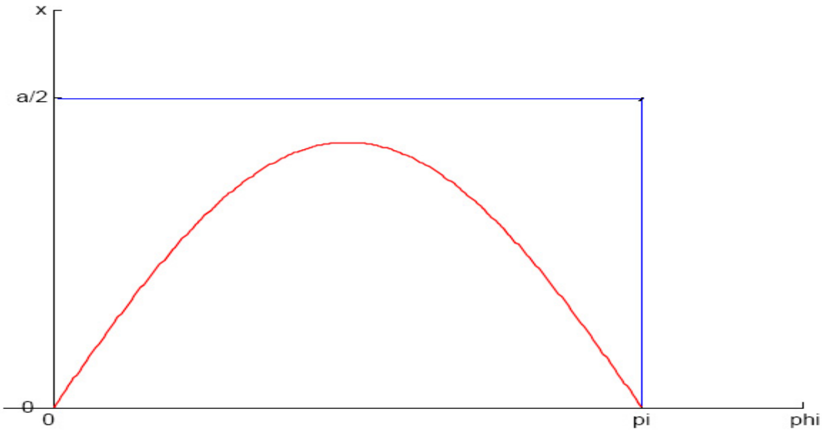


The position the needle lies in (the sample point) is decided by two parameters, the distance x between needle's midpoint and the line closest to it , and the angle φ between the needle and parallel lines.

The position the needle lies in (the sample point) is decided by two parameters, the distance x between needle's midpoint and the line closest to it , and the angle φ between the needle and parallel lines. So the sample space is

$$\Omega = \{(\varphi, x) \big| 0 \le \varphi \le \pi, 0 \le x \le \frac{a}{2}\}.$$

The needle crosses one of the parallel lines if and only if $x \le \frac{l}{2} \sin\varphi$ (denote this area by g).


x
a/2
0
0
pi
phi

Hence the desired probability is

$$P = \frac{\text{the area of } g}{\text{the area of } \Omega}$$

Hence the desired probability is

$$\begin{aligned} P &= \frac{\text{the area of } g}{\text{the area of } \Omega} \\ &= \frac{\int_0^\pi \frac{l}{2}\sin\varphi d\varphi}{\pi a/2} = \frac{2l}{a\pi}. \end{aligned}$$

Monte Carlo method If we know the value of P, then we can obtain

$$\pi = \frac{2l}{a} P.$$

Since

$$\text{probability}\, P \approx \text{frequency } \frac{n}{N},$$

the latter can be obtained from a large number of repeated independent experiments. Then

$$\pi \approx \frac{2l}{a} \frac{n}{N}.$$

In history, one of the best approximate values is

$\pi \approx 3.1415929.$

In history, one of the best approximate values is

$\pi \approx 3.1415929.$

Nowadays, repeated independent experiments can be simulated by computator. And so the probability can by approximated by simulation.

In history, one of the best approximate values is

$\pi \approx 3.1415929$.

Nowadays, repeated independent experiments can be simulated by computator. And so the probability can by approximated by simulation.

Monte Carlo method is very popular in computation.

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;
2. Normalization(规范性): $P(\Omega) = 1$, where Ω is a sure event;

Properties of probability:

1. Non-negativity(非负性): $P(A) \geq 0$;
2. Normalization(规范性): $P(\Omega) = 1$, where Ω is a sure event;
3. Additivity(可加性): Suppose that A and B will never happen simultaneously and that $A + B$ stands for the event that A, or B, or both happen, then $P(A + B) = P(A) + P(B)$.

4. **Countable Additivity(可数可加性)**: Suppose that any two of A_i, $i = 1, 2, \cdots$, will never happen simultaneously (i.e., $A_i \cap A_j = \emptyset$, $i \neq j$) and that $\sum_{i=1}^{\infty} A_i$ stands for the event that at least one of A_i, $i = 1, 2, \cdots$, happens (i.e., $\bigcup_{i=1}^{\infty} A_i$), then

$$P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$