Statistical definition of probability:

$$F_N(A) = \frac{n_A}{N} \text{ stabilizes to } P(A).$$

Statistical definition of probability:

$$F_N(A) = \frac{n_A}{N} \text{ stabilizes to } P(A).$$

Classical probability:

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

Statistical definition of probability:

$$F_N(A) = \frac{n_A}{N} \text{ stabilizes to } P(A).$$

Classical probability:

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

Geometrical probability:

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

1.3 The axiomatic definition of probability

(概率的公理化定义)

Andrey Nikolaevich Kolmogorov (April 1903–October 1987)

Events

A sample point (often called an elementary outcome) is an event that cannot be broken down into some combination of other events. In a random experiment, besides elementary outcomes——sample points, we are interested in some other results.

Example

A bag contains 10 balls, 3 of which are red, 3 white and 4 black, red–$1, 2, 3$, white– $4, 5, 6$, black– $7, 8, 9, 10$. If a ball is drawn at random, then the sample space is

Example

A bag contains 10 balls, 3 of which are red, 3 white and 4 black, red–$1, 2, 3$, white– $4, 5, 6$, black– $7, 8, 9, 10$. If a ball is drawn at random, then the sample space is

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

Consider the following results:

$$A = \{\text{the ball drawn out is red or white}\};$$

$$B = \{\text{the number of the balls drawn out is less than } 5\};$$

$$C = \{\text{the ball drawn out is not a red one}\}.$$

These are all events.

Consider the following results:

$$A = \{\text{the ball drawn out is red or white}\};$$

$$B = \{\text{the number of the balls drawn out is less than } 5\};$$

$$C = \{\text{the ball drawn out is not a red one}\}.$$

These are all events.

$$A = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}; \quad B = \{\omega_1, \omega_2, \omega_3, \omega_4\};$$

$$C = \{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\}.$$

An event A is defined as a certain subset of the sample space Ω, a certain set composed of sample points.

An event A is defined as a certain subset of the sample space Ω, a certain set composed of sample points.

Ω——-sure event

$\emptyset$——- impossible event

An event A is defined as a certain subset of the sample space Ω, a certain set composed of sample points.

Ω—-sure event

$\emptyset$—- impossible event

If the outcome $\omega \in A$, the we call the event A happens.

Some relations of Events:

- $B \subset A$—If event A happens whenever event B happens
- union: $A \cup B$
- intersection (or product): $A \cap B$ or AB
- complement: A^c (or $\overline{A}$)
- mutually exclusive (or disjoint): $A \cap B = \emptyset$

- If $A \cap B = \emptyset$, then we define $A + B = A \cup B$
- $A - B = A\overline{B}$— the event that A happens but B does not happen

Commutative law (交换律) $A \cup B = B \cup A, AB = BA$;

Associative law (结合律)

$$(A \cup B) \cup C = A \cup (B \cup C), (AB)C = A(BC);$$

Distributive law (分配律)

$$(A \cup B) \cap C = AC \cup BC, (A \cap B) \cup C = (A \cup C) \cap (B \cup C);$$

de Morgoan's law

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \qquad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

Example

Suppose that A, B, C are three events, then

Example

Suppose that A, B, C are three events, then

(1) {both A and B come up while C does not}$= AB\overline{C} = AB - C = AB - ABC$;

Example

Suppose that A, B, C are three events, then

(1) {both A and B come up while C does not}$= AB\overline{C} = AB - C = AB - ABC$;

(2) $\{A, B$ and C all come up$\}$ can be written as ABC;

Example

Suppose that A, B, C are three events, then

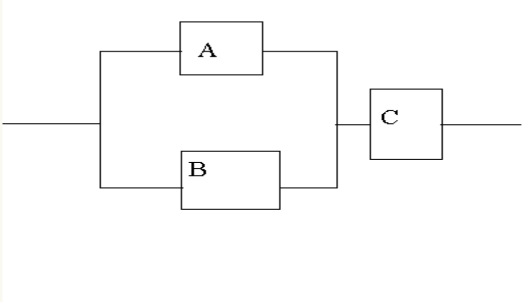
(1) {both A and B come up while C does not}$= AB\overline{C} = AB - C = AB - ABC$;

(2) $\{A, B$ and C all come up$\}$ can be written as ABC;

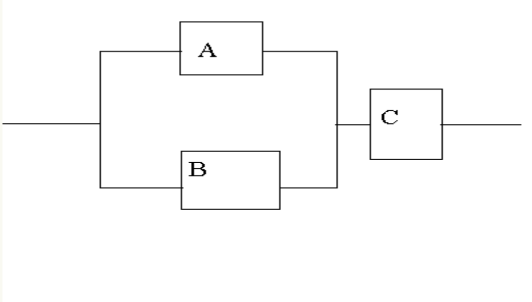
(3) {at least one of A, B and C come up} can be written as $A \cup B \cup C$ or

$$A\overline{B}\,\overline{C} + \overline{A}B\overline{C} + \overline{A}\,\overline{B}C + AB\overline{C} + \overline{A}BC + A\overline{B}C + ABC.$$

Example



Example



$$\{\text{the system works orderly}\} = (A \cup B)C \text{ or } AC \cup BC.$$

Probability space A probability space contains three basic elements $(\Omega, \mathcal{F}, P)$

Ω—the sample space

$\mathcal{F}$–σ-fields, σ-algebra—the family of events

P–probability

Geometrical probability: $A_g = \{$ a sample point falls into region $g \subset \Omega\}$,

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

Geometrical probability: $A_g = \{$ a sample point falls into region $g \subset \Omega\}$,

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

If the measure of g does not exists, we can not define the probability of A_g.

Geometrical probability: $A_g = \{$ a sample point falls into region $g \subset \Omega\}$,

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

If the measure of g does not exists, we can not define the probability of A_g. So, we do not take such A_g as an event.

The σ-algebra $\mathcal{F}$ of events is a family of Ω subsets satisfying:

(1) $\Omega \in \mathcal{F}$;

(2) If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$;

(3) If $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

$\mathcal{F}$ satisfying the above three hypotheses is termed a σ-algebra (or σ-field) in Ω and the elements of $\mathcal{F}$ (subsets of Ω) are called events.

Properties about σ-algebra of events:

Properties about σ-algebra of events: (4) $\emptyset \in \mathcal{F}$ $(\emptyset = \overline{\Omega})$;

Properties about σ-algebra of events: (4) $\emptyset \in \mathcal{F}$ ($\emptyset = \overline{\Omega}$);

(5) If $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{F}$, then $\cap_{n=1}^{\infty} A_n \in \mathcal{F}$. Indeed,

$$\bigcap_{n=1}^{\infty} A_n = \overline{\bigcup_{n=1}^{\infty} \overline{A_n}};$$

Properties about σ-algebra of events: (4) $\emptyset \in \mathcal{F}$ ($\emptyset = \overline{\Omega}$);

(5) If $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{F}$, then $\cap_{n=1}^{\infty} A_n \in \mathcal{F}$. Indeed,

$$\bigcap_{n=1}^{\infty} A_n = \overline{\bigcup_{n=1}^{\infty} \overline{A_n}};$$

(6) If $A_1, A_2, \cdots, A_n \in \mathcal{F}$, then $\cup_{k=1}^{n} A_k \in \mathcal{F}$, $\cap_{k=1}^{n} A_k \in \mathcal{F}$.

We conclude from the above statement that inevitable event, the impossible event, and complements, finite unions, finite intersections, countable unions, countable intersections of events are all still events and thus the operations like complement, union and intersection in σ-algebra of events are all meaningful.

Examples:

- $\mathcal{F}_1 = \{\emptyset, \Omega\}$

Examples:

- $\mathcal{F}_1 = \{\emptyset, \Omega\}$

Examples:

- $\mathcal{F}_1 = \{\emptyset, \Omega\} \quad \mathcal{F}_2 = \{$ all subset of $\Omega\}$
- $\Omega = \{\omega_1, \omega_2, \cdots\}$—$\mathcal{F} = \{$ all subset of $\Omega\}$

Examples:

- $\mathcal{F}_1 = \{\emptyset, \Omega\}$ $\mathcal{F}_2 = \{$ all subset of $\Omega\}$
- $\Omega = \{\omega_1, \omega_2, \cdots\}$—$\mathcal{F} = \{$ all subset of $\Omega\}$
- If $\Omega = \boldsymbol{R}^1$, the family of all intervals and sets of the unions, intersections and complements of them is chosen to be $\mathcal{F}$—-one-dimensional Borel σ-algebra, denoted by $\mathscr{B}$.

Examples:

- $\mathcal{F}_1 = \{\emptyset, \Omega\}$ $\quad \mathcal{F}_2 = \{$ all subset of $\Omega\}$
- $\Omega = \{\omega_1, \omega_2, \cdots\}$—$\mathcal{F} = \{$ all subset of $\Omega\}$
- If $\Omega = \boldsymbol{R}^1$, the family of all intervals and sets of the unions, intersections and complements of them is chosen to be $\mathcal{F}$—-one-dimensional Borel σ-algebra, denoted by $\mathscr{B}$.
- If $\Omega = \boldsymbol{R}^n$, $\mathscr{B}^n$ – n-dimensional Borel σ-algebra

- 关于事件域的取法

- 关于事件域的取法

- 关于事件域的取法

◇ 如果$\Omega = \{\omega_1, \omega_2, \ldots\}$只含有限个或者可列个点，这时通常取

$$\mathcal{F} = \{A : A \subset \Omega\}.$$

- 关于事件域的取法

◇ 如果$\Omega = \{\omega_1, \omega_2, \dots\}$只含有限个或者可列个点，这时通常取

$$\mathcal{F} = \{A : A \subset \Omega\}.$$

◇ 如果Ω为实数集R(或者R中的一个区间$[a, b]$)，这时R的一些子集上无法定义合理的概率. 一般取$\mathcal{F}$为$R([a, b])$上的Borel集类:

$$\mathscr{B} = \sigma\Big(\{(a, b] : a < b\}\Big).$$

$\mathscr{B}$中的元素称为Borel集合.

设$\mathcal{C}$为一集合类，$\sigma(\mathcal{C})$表示包含$\mathcal{C}$的最小σ-域.

即

- $\sigma(\mathcal{C})$是σ-域，
- $\mathcal{C} \subset \sigma(\mathcal{C})$，
- 并且，如果$\mathcal{L}$也是包含$\mathcal{C}$的σ-域，则必有$\sigma(\mathcal{C}) \subset \mathcal{L}$.

Theorem

设$\mathcal{C}$为Ω上的一个集合类，那么存在唯一的一个σ-域$\mathcal{G}$ 包含$\mathcal{C}$，并且对任何包含$\mathcal{C}$的σ-域$\mathcal{L}$ 有，$\mathcal{G} \subset \mathcal{L}$. 即，$\sigma(\mathcal{C})$存在且唯一.

证明：记

$$\mathscr{L} = \left\{ \mathcal{L} : \mathcal{L} \supset \mathcal{C}, \mathcal{L} \text{ 为}\sigma\text{-域} \right\}.$$

则$2^\Omega \in \mathcal{L}$. 所以$\mathscr{L}$为空.

证明：记

$$\mathscr{L} = \left\{ \mathcal{L} : \mathcal{L} \supset \mathcal{C}, \mathcal{L} \text{ 为}\sigma\text{-域} \right\}.$$

则$2^\Omega \in \mathcal{L}$. 所以$\mathscr{L}$为空. 令

$$\mathcal{G} = \bigcap_{\mathcal{L} \in \mathscr{L}} \mathcal{L}.$$

显然$\mathcal{G} \supset \mathcal{C}$. 下面验证$\mathcal{G}$即为所求.

证明：记

$$\mathscr{L} = \Big\{\mathcal{L} : \mathcal{L} \supset \mathcal{C}, \mathcal{L}\ 为\sigma\text{-域}\Big\}.$$

则$2^\Omega \in \mathcal{L}$. 所以$\mathscr{L}$为空. 令

$$\mathcal{G} = \bigcap_{\mathcal{L}\in\mathscr{L}} \mathcal{L}.$$

显然$\mathcal{G} \supset \mathcal{C}$. 下面验证$\mathcal{G}$即为所求. 首先, $\mathcal{G}$为σ域. 事实上,
(1) 因为Ω属于每个$\mathcal{L}$ (这是因为$\mathscr{L}$的每个元素$\mathcal{L}$为σ-域), 所以$\Omega \in \mathcal{G}$.

(2) 若$A \in \mathcal{G}$, 则对每一个$\mathcal{L} \in \mathscr{L}$, $A \in \mathcal{L}$. 所以$\overline{A} \in \mathcal{L}$ (这是因为$\mathcal{L}$为σ-域), 所以

$$\overline{A} \in \bigcap_{\mathcal{L} \in \mathscr{L}} \mathcal{L} = \mathcal{G}.$$

(2) 若$A\in\mathcal{G}$, 则对每一个$\mathcal{L}\in\mathscr{L}$, $A\in\mathcal{L}$. 所以$\overline{A}\in\mathcal{L}$ (这是因为$\mathcal{L}$为σ-域), 所以

$$\overline{A}\in\bigcap_{\mathcal{L}\in\mathscr{L}}\mathcal{L}=\mathcal{G}.$$

(3) 若$A_i\in\mathcal{G}$, 则对每一个$\mathcal{L}\in\mathscr{L}$, $A_i\in\mathcal{L}$. 所以$\bigcup_{i=1}^{\infty}A_i\in\mathcal{L}$ (这是因为$\mathcal{L}$为σ-域), 所以

$$\bigcup_{i=1}^{\infty}A_i\in\bigcap_{\mathcal{L}\in\mathscr{L}}\mathcal{L}=\mathcal{G}.$$

最后，如果$\mathcal{L}$为包含$\mathcal{C}$的σ-域，那么$\mathcal{L} \in \mathscr{L}$，所以$\mathcal{G} \subset \mathcal{L}$.

命题 若$\Omega = \{\omega_1, \omega_2, \ldots,\}$. $\mathcal{C} = \{\{\omega_1\}, \{\omega_2\}, \ldots,\}$. 则

$$\sigma(\mathcal{C}) = 2^{\Omega}.$$

命题 若$R=(-\infty,\infty)$表示数直线，则下列集类生成相同的σ域：

1. $\{(a,b]:a,b\in R\}$;
2. $\{(a,b):a,b\in R\}$;
3. $\{[a,b]:a,b\in R\}$;
4. $\{(-\infty,b]:a,b\in R\}$;
5. $\{(r_1,r_2):r_1,r_2$为有理数$\}$;
6. $\{G:G$为R中的开集$\}$;
7. $\{F:F$为R中的闭集$\}$.

证明： 由于

$$(a,b]=\bigcap_n(a,b+1/n),\quad (a,b)=\bigcup_n(a,b-1/n],$$

所以(1), (2)中集类生成相同的σ域. 同样,(1), (3)中集类生成相同的σ域.

证明：由于

$$(a,b]=\bigcap_n (a,b+1/n), \quad (a,b)=\bigcup_n (a,b-1/n],$$

所以(1), (2)中集类生成相同的σ域. 同样,(1), (3)中集类生成相同的σ域. 此外，由

$$(-\infty,b]=\bigcup_n (-n,b], \quad (a,b]=(-\infty,b]\setminus(-\infty,a],$$

(1), (4)中集类生成相同的σ域.

由

$$(a,b)=\bigcup_{a<r_1<r_2<b}(r_1,r_2)$$

可知(2), (5)中集类生成相同的σ域.

由

$$(a,b)=\bigcup_{a<r_1<r_2<b}(r_1,r_2)$$

可知(2), (5)中集类生成相同的σ域. 由于R中任一开集可以表示为至多可列个开区间的并, 故(2), (6)中集类生成相同的σ域.

由

$$(a,b)=\bigcup_{a<r_1<r_2<b}(r_1,r_2)$$

可知(2), (5)中集类生成相同的σ域. 由于R中任一开集可以表示为至多可列个开区间的并, 故(2), (6)中集类生成相同的σ域. 最后,利用开集的余集为闭集, 可得(6), (7)中集类生成相同的σ域.

Definition

Probability

Definition

Probability P is a real function defined on $\mathcal{F}$: $A \longrightarrow P(A)$, satisfying

- P_1 (non-negativity) $P(A) \geq 0$ for all $A \in \mathcal{F}$;

Definition

Probability P is a real function defined on $\mathcal{F}$: $A \longrightarrow P(A)$, satisfying

- P_1 (non-negativity) $P(A) \geq 0$ for all $A \in \mathcal{F}$;
- P_2 (normalization condition) $P(\Omega) = 1$;

Definition

Probability P is a real function defined on $\mathcal{F}$: $A \longrightarrow P(A)$, satisfying

- P_1 (non-negativity) $P(A) \geq 0$ for all $A \in \mathcal{F}$;
- P_2 (normalization condition) $P(\Omega) = 1$;
- P_3 (countable additivity) If $A_1, \cdots, A_n, \cdots$ are mutually disjoint events ($A_i A_j = \emptyset$, $i \neq j$), then

$$P(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

Remark

A real function defined m on $\mathcal{F}$ satisfying P_1 and P_3 is called a measure.

Example

We shall construct a probability measure P on an arbitrary σ-field $\mathcal{F}$ in an arbitrary non-empty Ω. Suppose that $\omega_0 \in \Omega$, define

$$P(A) = I_A(\omega_0) = \begin{cases} 1, & \omega_0 \in A, \\ 0, & \text{otherwise} \end{cases}$$

for $A \in \mathcal{F}$.

P is clearly a discrete probability measure— a unit mass at ω_0.

Example

$\Omega = \{\omega_1, \omega_2, \ldots\}$. Suppose $p_j = P(\{\omega_j\})$, and let $\mathcal{F} = \{$ all subsets of $\Omega\}$. Define

$$P(A) = \sum_{\omega_j \in A} p_j.$$

The $(\Omega, \mathcal{F}, P)$ is a probability space.

Example

Let $\Omega = (0,1]$. $\mathcal{F} = \mathscr{B}_{(0,1]}$ be Borel field on $(0,1]$, and m be the Lebesgue measure. The $(\Omega, \mathcal{F}, m)$ is a probability space.

Remark

There is an unique measure m on $(\boldsymbol{R}, \mathscr{B})$ satisfying

$$m\big((a,b]\big) = b - a, \quad \forall a < b.$$

This measure is called the Lebesgue measure on $\boldsymbol{R}$.

Remark

In general, there is an unique measure m on $(\boldsymbol{R}^n, \mathscr{B}^n)$ satisfying

$$m\big((a_1, b_1] \times \cdots \times (a_n, b_n]\big) = (b_1 - a_1) \cdots (b_n - a_n),$$
$$\forall a_i < b_i.$$

It is called the Lebesgue measure on $\boldsymbol{R}^n$.

Example

Let $\Omega = (-\infty, +\infty)$. $\mathcal{F} = \mathscr{B}_{(-\infty,\infty)}$ be Borel field on $(-\infty, \infty)$, and m be the Lebesgue measure. $p(x) \geq 0$ is a function with $\int_{-\infty}^{\infty} p(x)dx = 1$. Define

$$P(A) = \int_A p(x)dx.$$

Then $(\Omega, \mathcal{F}, P)$ is a probability space.

无特别声明时，以后的叙述都建立在概率空间的基础上，所遇到的Ω的子集都假定为事件.

Properties for probability:

1. $P(\emptyset) = 0.$

Properties for probability:

1. $P(\emptyset) = 0.$

Properties for probability:

1. $P(\emptyset) = 0$.

   Proof.

$$\begin{aligned} P(\Omega) &= P(\Omega + \emptyset + \emptyset + \cdots) \\ &= P(\Omega) + P(\emptyset) + P(\emptyset) + \cdots \end{aligned}$$

   implies $P(\emptyset) = 0$.

2. (*finite additivity*) If $A_iA_j = \emptyset$ for $i \neq j$, then $P(\sum_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.

2. (*finite additivity*) If $A_iA_j = \emptyset$ for $i \neq j$, then $P(\sum_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.

Proof.

$$\begin{aligned} P(\sum_{i=1}^n A_i) &= P(A_1 + \cdots + A_n + \emptyset + \cdots) \\ &= \sum_{i=1}^n P(A_i) + \sum_{i=n+1}^\infty P(\emptyset) \\ &= \sum_{i=1}^n P(A_i). \end{aligned}$$

3. $P(\overline{A}) = 1 - P(A).$

3. $P(\overline{A}) = 1 - P(A)$.

Proof. $P(A) + P(\bar{A}) = P(\Omega) = 1$.

4. If $B \subset A$, then $P(A-B) = P(A) - P(B)$.

4. If $B \subset A$, then $P(A - B) = P(A) - P(B)$.

4. If $B \subset A$, then $P(A-B) = P(A) - P(B)$.

Proof. The equality is from $A = B + (A-B)$.

4. If $B \subset A$, then $P(A-B) = P(A) - P(B)$.

Proof. The equality is from $A = B + (A-B)$.

Corollary

(a) (Monotonicity) If $B \subset A$, *then*

$$P(B) \leq P(A).$$

(b) $0 \leq P(A) \leq 1$.

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

Proof. The equality is from

$$A \cup B = A + (B - AB).$$

and $AB \subset B$.

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

Proof. The equality is from

$$A \cup B = A + (B - AB).$$

and $AB \subset B$.

6. $P(A - B) = P(A) - P(AB)$

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

   Proof. The equality is from

   $$A \cup B = A + (B - AB).$$

   and $AB \subset B$.

6. $P(A - B) = P(A) - P(AB)$

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

Proof. The equality is from

$$A \cup B = A + (B - AB).$$

and $AB \subset B$.

6. $P(A - B) = P(A) - P(AB)$

Proof. Note that $A - B = A - AB$ and $AB \subset A$.

7. (Exclusion-inclusion) (Jordan 公式)

$$
\begin{aligned}
& P(A_1 \cup A_2 \cup \cdots \cup A_n) \\
= & \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \cdots \\
& + (-1)^{r-1} \sum_{i_1 < i_2 < \cdots < i_r} P(A_{i_1} A_{i_2} \cdots A_{i_r}) \\
& + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n).
\end{aligned}
$$

Proof. By (5) and the induction.

Proof. By (5) and the induction.

8. (sub-additivity) (Boole's inequality)

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Proof. By (5) and the induction.

8. (sub-additivity) (Boole's inequality)

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Proof. By (5) and the induction.

8. (sub-additivity) (Boole's inequality)

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Proof. Let $B_1 = A_1$, $B_2 = A_2\overline{A}_1$, $\cdots$, $B_n = A_n\overline{A}_1\cdots\overline{A}_{n-1}$, $\cdots$.

Proof. By (5) and the induction.

8. (sub-additivity) (Boole's inequality)

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Proof. Let $B_1 = A_1$, $B_2 = A_2\overline{A}_1$, $\cdots$, $B_n = A_n\overline{A}_1 \cdots \overline{A}_{n-1}$, $\cdots$.
Then $B_i \cap B_j = \emptyset$ $(i \neq j)$, and $B_i \subset A_i$. Also $\bigcup_{i=1}^{\infty} A_i = \sum_{i=1}^{\infty} B_i$.

Proof. By (5) and the induction.

8. (sub-additivity) (Boole's inequality)

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Proof. Let $B_1 = A_1$, $B_2 = A_2\overline{A}_1$, $\cdots$, $B_n = A_n\overline{A}_1 \cdots \overline{A}_{n-1}$, $\cdots$.
Then $B_i \cap B_j = \emptyset$ $(i \neq j)$, and $B_i \subset A_i$. Also $\bigcup_{i=1}^{\infty} A_i = \sum_{i=1}^{\infty} B_i$.
In fact, it is obvious that $\bigcup_{i=1}^{\infty} A_i \supset \sum_{i=1}^{\infty} B_i$.

On the other hand, suppose $\omega \in \bigcup_{i=1}^{\infty} A_i$. Then $\omega \in A_i$ for some i.

On the other hand, suppose $\omega \in \bigcup_{i=1}^{\infty} A_i$. Then $\omega \in A_i$ for some i. Let n be the smallest one of such i's. Then $\omega \in A_n$, but $\omega \notin A_j$, $j = 1, \cdots, n-1$. So $\omega \in B_n$. It follows that $\bigcup_{i=1}^{\infty} A_i \subset \sum_{i=1}^{\infty} B_i$.

On the other hand, suppose $\omega \in \bigcup_{i=1}^{\infty} A_i$. Then $\omega \in A_i$ for some i. Let n be the smallest one of such i's. Then $\omega \in A_n$, but $\omega \notin A_j$, $j = 1, \cdots, n-1$. So $\omega \in B_n$. It follows that $\bigcup_{i=1}^{\infty} A_i \subset \sum_{i=1}^{\infty} B_i$. Now, from the countable additivity and the monotonicity,

$$P(\bigcup_{i=1}^{\infty} A_i) = P(\sum_{i=1}^{\infty} B_i)$$

On the other hand, suppose $\omega \in \bigcup_{i=1}^{\infty} A_i$. Then $\omega \in A_i$ for some i. Let n be the smallest one of such i's. Then $\omega \in A_n$, but $\omega \notin A_j$, $j = 1, \cdots, n-1$. So $\omega \in B_n$. It follows that $\bigcup_{i=1}^{\infty} A_i \subset \sum_{i=1}^{\infty} B_i$. Now, from the countable additivity and the monotonicity,

$$P(\bigcup_{i=1}^{\infty} A_i) = P(\sum_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i)$$

On the other hand, suppose $\omega \in \bigcup_{i=1}^{\infty} A_i$. Then $\omega \in A_i$ for some i. Let n be the smallest one of such i's. Then $\omega \in A_n$, but $\omega \notin A_j$, $j = 1, \cdots, n-1$. So $\omega \in B_n$. It follows that $\bigcup_{i=1}^{\infty} A_i \subset \sum_{i=1}^{\infty} B_i$. Now, from the countable additivity and the monotonicity,

$$P(\bigcup_{i=1}^{\infty} A_i) = P(\sum_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Example 4. A bag contains n $(n \geq 3)$ balls numbered $1, 2, \cdots, n$ respectively. Take three balls randomly; find the probability that at least one of ball 1 and ball 2 is taken.

Solution (1).

Solution (1). Let A_i={ ball i is taken}, $i = 1, 2$, then the desired probability is

Solution (1). Let A_i=$\{$ ball i is taken$\}$, $i = 1, 2$, then the desired probability is

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1A_2)$$
$$= \frac{\binom{n-1}{2}}{\binom{n}{3}} + \frac{\binom{n-1}{2}}{\binom{n}{3}} - \frac{\binom{n-2}{1}}{\binom{n}{3}}.$$

Solution (1). Let A_i={ ball i is taken}, $i = 1, 2$, then the desired probability is

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1A_2)$$
$$= \frac{\binom{n-1}{2}}{\binom{n}{3}} + \frac{\binom{n-1}{2}}{\binom{n}{3}} - \frac{\binom{n-2}{1}}{\binom{n}{3}}.$$

Solution (2). The complement of $A_1 \cup A_2$ is $\overline{A_1 \cup A_2} = \overline{A_1} \cap \overline{A_2}$.

Solution (1). Let A_i={ ball i is taken}, $i = 1, 2$, then the desired probability is

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1A_2)$$
$$= \frac{\binom{n-1}{2}}{\binom{n}{3}} + \frac{\binom{n-1}{2}}{\binom{n}{3}} - \frac{\binom{n-2}{1}}{\binom{n}{3}}.$$

Solution (2). The complement of $A_1 \cup A_2$ is $\overline{A_1 \cup A_2} = \overline{A_1} \cap \overline{A_2}$. So

$$P(A_1 \cup A_2) = 1 - P(\overline{A_1} \cap \overline{A_2})$$

Solution (1). Let A_i={ ball i is taken}, $i = 1, 2$, then the desired probability is

$$\begin{aligned} P(A_1 \cup A_2) =& P(A_1) + P(A_2) - P(A_1A_2) \\ =& \frac{\binom{n-1}{2}}{\binom{n}{3}} + \frac{\binom{n-1}{2}}{\binom{n}{3}} - \frac{\binom{n-2}{1}}{\binom{n}{3}}. \end{aligned}$$

Solution (2). The complement of $A_1 \cup A_2$ is $\overline{A_1 \cup A_2} = \overline{A_1} \cap \overline{A_2}$. So

$$P(A_1 \cup A_2) = 1 - P(\overline{A_1} \cap \overline{A_2}) = 1 - \frac{\binom{n-2}{3}}{\binom{n}{3}}.$$

Example 5: (Match problem) Suppose that a person types n letters, types the corresponding addresses on n envelopes, and then places the n letters in the n envelopes in random manner. Find the probability p_n that at least one letter will be placed in the correct envelope.

Solution. Let A_i={letter i is placed in the correct envelope}, $i = 1, 2, \cdots, n$, then the required probability p_n is $P(\cup_{i=1}^n A_i)$.

Solution. Let A_i={letter i is placed in the correct envelope}, $i = 1, 2, \cdots, n$, then the required probability p_n is $P(\cup_{i=1}^{n} A_i)$. Note that

$$P(A_i) = \frac{1}{n};$$

Solution. Let A_i={letter i is placed in the correct envelope}, $i = 1, 2, \cdots, n$, then the required probability p_n is $P(\cup_{i=1}^n A_i)$. Note that

$$P(A_i) = \frac{1}{n};$$

$$P(A_iA_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)};\ 1 \le i < j \le n,$$

Solution. Let A_i={letter i is placed in the correct envelope}, $i = 1, 2, \cdots, n$, then the required probability p_n is $P(\cup_{i=1}^n A_i)$. Note that

$$P(A_i) = \frac{1}{n};$$

$$P(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)};\ 1 \le i < j \le n,$$

$$\cdots, P(A_{i_1} A_{i_2} \cdots A_{i_k}) = \frac{(n-k)!}{n!};\ 1 \le i_1 < \cdots i_k \le n,$$

Solution. Let A_i={letter i is placed in the correct envelope}, $i = 1, 2, \cdots, n$, then the required probability p_n is $P(\cup_{i=1}^n A_i)$. Note that

$$P(A_i) = \frac{1}{n};$$

$$P(A_iA_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)};\ 1 \le i < j \le n,$$

$$\cdots, P(A_{i_1}A_{i_2}\cdots A_{i_k}) = \frac{(n-k)!}{n!};\ 1 \le i_1 < \cdots i_k \le n,$$

$$\cdots, P(A_1A_2\cdots A_n) = \frac{1}{n!}.$$

Hence

Hence

$$\begin{aligned} p_n &= n \cdot \frac{1}{n} - \binom{n}{2} \frac{1}{n(n-1)} \\ &\quad + \binom{n}{3} \frac{1}{n(n-1)(n-2)} + \cdots \\ &\quad + (-1)^{k-1} \binom{n}{k} \frac{1}{n(n-1)\cdots(n-k+1)} \\ &\quad + \cdots + (-1)^{n-1} \frac{1}{n!} \end{aligned}$$

Hence

$$
\begin{aligned}
p_n &= n\cdot\frac{1}{n}-\binom{n}{2}\frac{1}{n(n-1)} \\
&\quad +\binom{n}{3}\frac{1}{n(n-1)(n-2)}+\cdots \\
&\quad +(-1)^{k-1}\binom{n}{k}\frac{1}{n(n-1)\cdots(n-k+1)} \\
&\quad +\cdots+(-1)^{n-1}\frac{1}{n!} \\
&= 1-\frac{1}{2!}+\frac{1}{3!}-\cdots+(-1)^{n-1}\frac{1}{n!}.
\end{aligned}
$$

$$p_\infty = 1 - 1/e.$$

The probability in Matching Problem

n	5	6	7	9	∞
p_n	0.633333	0.631944	0.632143	0.632121	0.632121

Example

(**涂色问题**) 平面上的n个点和连接各点之间的连线叫做一个完全图,记作G. 点称作顶点, 顶点之间的连线叫做边,共有$\binom{n}{2}$条边. 给定一个整数k. G中任意k个顶点连同相应的边也构成一个k个顶点的完全子图, G中共有$\binom{n}{k}$个这样的子图, 记作G_i,$i = 1, \cdots, \binom{n}{k}$. 现将图$G$的每条边图成红色或蓝色. 问是否有一种涂色方法, 使得没有一个子图G_i的$\binom{k}{2}$条边同一颜色.

Example

(**涂色问题**) 平面上的n个点和连接各点之间的连线叫做一个完全图,记作G. 点称作顶点，顶点之间的连线叫做边,共有$\binom{n}{2}$条边. 给定一个整数k. G中任意k个顶点连同相应的边也构成一个k个顶点的完全子图，G中共有$\binom{n}{k}$个这样的子图，记作G_i,$i=1,\cdots,\binom{n}{k}$. 现将图G的每条边图成红色或蓝色. 问是否有一种涂色方法，使得没有一个子图G_i的$\binom{k}{2}$条边同一颜色.

这是一个确定性的问题，下面我们用概率方法(probabilistic method)证明，当n不太大时(相对于k)，答案是肯定的.

我们对G的边进行随机涂色，每条边为红色和蓝色的概率均为1/2. 记事件

$$E_i = \{\text{子图}G_i\text{ 各边的颜色相同}\}.$$

那么$\bigcup_i E_i$就表示至少存在一个子图使得它的各边颜色相同.

我们对G的边进行随机涂色, 每条边为红色和蓝色的概率均为$1/2$. 记事件

$$E_i=\{子图G_i\text{ 各边的颜色相同}\}.$$

那么$\bigcup_i E_i$就表示至少存在一个子图使得它的各边颜色相同.
易知

$$\begin{aligned}P(E_i)=&P(G_i\text{ 各边的均为红色})+P(G_i\text{ 各边的均为蓝色})\\=&2\frac{1}{2^{\binom{k}{2}}}=\left(\frac{1}{2}\right)^{k(k-1)/2-1}.\end{aligned}$$

从而

$$
\begin{aligned}
P(\bigcup_i E_i) &\le \sum_i P(E_i) \\
&= \binom{n}{k}\left(\frac{1}{2}\right)^{k(k-1)/2-1}.
\end{aligned}
$$

从而

$$
\begin{aligned}
P(\bigcup_i E_i) &\le \sum_i P(E_i) \\
&= \binom{n}{k}\left(\frac{1}{2}\right)^{k(k-1)/2-1}.
\end{aligned}
$$

所以当

$$
\binom{n}{k} < 2^{k(k-1)/2-1},
$$

时，$P(\bigcup_i E_i) < 1$.

从而

$$P(\bigcup_i E_i) \le \sum_i P(E_i) = \binom{n}{k}\left(\frac{1}{2}\right)^{k(k-1)/2-1}.$$

所以当

$$\binom{n}{k} < 2^{k(k-1)/2-1},$$

时，$P(\bigcup_i E_i) < 1$. 这说明，$\bigcap_i \overline{E_i} \neq \emptyset$. 从而，在上述条件下，至少有一种涂色方法，使得没有一个有k个顶点、所有边同一颜色的完全子图G_i.

Remark

The method of introducing probability to a problem whose statement is purely deterministic has been called the probabilistic method.

Probabilistic method 为解决数学中的一些困难问题带来很多方便.

例如: 要找到连续而又处处不可微的函数不是十分显然的. 而利用概率的方法, 可以证明这样的函数比可微函数多得多.

Continuity of probability measure

$(\Omega, \mathcal{F}, P)$ —a probability space.

Suppose $A_1, A_2, \cdots$, is a sequence of increasing events, i.e., $A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$. Denote

$$A = \cup_{n=1}^{\infty} A_n \triangleq \lim_{n\to\infty} A_n.$$

Theorem

Suppose that $A_1, A_2, \cdots,$ is a sequence of increasing events with A as its limit, then

$$P(A) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

Proof. Let $A_0 = \emptyset$, $B_k = A_k - A_{k-1}, k = 1, 2, \cdots$, then

Proof. Let $A_0 = \emptyset$, $B_k = A_k - A_{k-1}, k = 1, 2, \cdots$, then

$$A = B_1 \cup B_2 \cup B_3 \cup \cdots ,$$

the union of a series of disjoint events.

Proof. Let $A_0 = \emptyset$, $B_k = A_k - A_{k-1}, k = 1, 2, \cdots$, then

$$A = B_1 \cup B_2 \cup B_3 \cup \cdots ,$$

the union of a series of disjoint events. By the countable additivity of probability, we have

$$\begin{aligned} P(A) =& P(A_1) + P(B_2) + P(B_3) + \cdots \\ =& \lim_{n\to\infty} \sum_{k=1}^{n} P(B_k). \end{aligned}$$

Since $B_1 + \cdots + B_n + \emptyset + \cdots = A_n$, by the the countable additivity of probability again we have

Since $B_1 + \cdots + B_n + \emptyset + \cdots = A_n$, by the the countable additivity of probability again we have

$$\sum_{k=1}^{n} P(B_k) + 0 + \cdots = P(A_n).$$

Since $B_1 + \cdots + B_n + \emptyset + \cdots = A_n$, by the the countable additivity of probability again we have

$$\sum_{k=1}^{n} P(B_k) + 0 + \cdots = P(A_n).$$

Therefore we have

$$P(A) = \lim_{n\to\infty} \sum_{k=1}^{n} P(B_k) = \lim_{n\to\infty} P(A_n),$$

Since $B_1 + \cdots + B_n + \emptyset + \cdots = A_n$, by the the countable additivity of probability again we have

$$\sum_{k=1}^{n} P(B_k) + 0 + \cdots = P(A_n).$$

Therefore we have

$$P(A) = \lim_{n\to\infty} \sum_{k=1}^{n} P(B_k) = \lim_{n\to\infty} P(A_n),$$

which completes the proof.

Similarly

Theorem

If A_n is a sequence of decreasing events and

$$A = \bigcap_{n=1}^{\infty} A_n \triangleq \lim_{n\to\infty} A_n,$$

then

$$P(A) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

In general, for a sequence of events $\{A_n\}$, we denote

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m, \ (A_n \text{ 至多有有限个不发生})$$

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m. \ (A_n \text{ 发生无穷多个})$$

In general, for a sequence of events $\{A_n\}$, we denote

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m, \ (A_n \text{ 至多有有限个不发生})$$

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m. \ (A_n \text{ 发生无穷多个})$$

It is easily seen that

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n.$$

If

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n,$$

we say that the limit of A_n exists and denote

$$\lim_{n\to\infty} A_n = \limsup_{n\to\infty} A_n.$$

It is easily seen that (Fatou Lemma)

$$\begin{aligned}P(\liminf_{n\to\infty} A_n) =& P(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m) \\ =& \lim_{n\to\infty} P(\bigcap_{m=n}^{\infty} A_m) \le \liminf_{n\to\infty} P(A_n).\end{aligned}$$

It is easily seen that (Fatou Lemma)

$$
\begin{aligned}
P(\liminf_{n\to\infty} A_n) =& P(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m) \\
=& \lim_{n\to\infty} P(\bigcap_{m=n}^{\infty} A_m) \le \liminf_{n\to\infty} P(A_n).
\end{aligned}
$$

$$
\begin{aligned}
P(\limsup_{n\to\infty} A_n) =& P(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m) \\
=& \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty} A_m) \ge \limsup_{n\to\infty} P(A_n).
\end{aligned}
$$

It is easily seen that (Fatou Lemma)

$$
\begin{aligned}
P(\liminf_{n\to\infty} A_n) =& P(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m) \\
=& \lim_{n\to\infty} P(\bigcap_{m=n}^{\infty} A_m) \le \liminf_{n\to\infty} P(A_n).
\end{aligned}
$$

$$
\begin{aligned}
P(\limsup_{n\to\infty} A_n) =& P(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m) \\
=& \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty} A_m) \ge \limsup_{n\to\infty} P(A_n).
\end{aligned}
$$

So, if $\lim_{n\to\infty} A_n$ exists, then $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$.

Theorem

For a sequence of events $\{A_n\}$, we have

$$P(\liminf_{n\to\infty} A_n) \le \liminf_{n\to\infty} P(A_n);$$

$$P(\limsup_{n\to\infty} A_n) \ge \limsup_{n\to\infty} P(A_n).$$

If $\lim_{n\to\infty} A_n$ exists, then

$$P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

Theorem

If $P:\mathcal{F}\to \boldsymbol{R}$ is a finite additive function on the σ-field $\mathcal{F}$, and if $A_n \searrow \emptyset$ for sets $A_n\in\mathcal{F}$ implies $P(A_n)\to 0$, then P is countably additive.

Proof. Suppose $B = \sum_{m=1}^{\infty} B_m$ for disjoint sets $B_m \in \mathcal{F}$.

Proof. Suppose $B = \sum_{m=1}^{\infty} B_m$ for disjoint sets $B_m \in \mathcal{F}$. Then $C_n = \sum_{m=n+1}^{\infty} B_m \in \mathcal{F}$, $B = \sum_{m=1}^{n} B_m + C_n$ and $C_n \searrow \emptyset$.

Proof. Suppose $B = \sum_{m=1}^{\infty} B_m$ for disjoint sets $B_m \in \mathcal{F}$. Then $C_n = \sum_{m=n+1}^{\infty} B_m \in \mathcal{F}$, $B = \sum_{m=1}^{n} B_m + C_n$ and $C_n \searrow \emptyset$. It follows that

$$P(B) = \sum_{m=1}^{n} P(B_m) + P(C_n)$$

due to the finite additivity, which gives

$P(B) - \sum_{m=1}^{n} P(B_m) = P(C_n) \to 0.$

Proof. Suppose $B = \sum_{m=1}^{\infty} B_m$ for disjoint sets $B_m \in \mathcal{F}$. Then $C_n = \sum_{m=n+1}^{\infty} B_m \in \mathcal{F}$, $B = \sum_{m=1}^{n} B_m + C_n$ and $C_n \searrow \emptyset$. It follows that

$$P(B) = \sum_{m=1}^{n} P(B_m) + P(C_n)$$

due to the finite additivity, which gives $P(B) - \sum_{m=1}^{n} P(B_m) = P(C_n) \to 0$. Hence

$$P(B) = \sum_{m=1}^{\infty} P(B_m).$$

Corollary

If $P : \mathcal{F} \to \boldsymbol{R}$ is a finite additive function on the σ-field $\mathcal{F}$, and if $A_n \nearrow \Omega$ for sets $A_n \in \mathcal{F}$ implies $P(A_n) \nearrow P(\Omega) < \infty$, then P is countably additive.

Example

Toss a fair coin infinitely many times independently, the probability that no head comes up is obviously 0. Use the above continuity theorem to explain this fact rigorously.

Proof: Let A_n is the event that no head comes up in the first n tosses. Then $A = \bigcap_{n=1}^{\infty} A_n$ is the event that no head comes up in all infinitely many tosses. Then we have $A_n \supset A_{n+1}$.

Proof: Let A_n is the event that no head comes up in the first n tosses. Then $A = \bigcap_{n=1}^{\infty} A_n$ is the event that no head comes up in all infinitely many tosses. Then we have $A_n \supset A_{n+1}$. So

$$P(A) = \lim_{n\to\infty} P(A_n) =$$

Proof: Let A_n is the event that no head comes up in the first n tosses. Then $A = \bigcap_{n=1}^{\infty} A_n$ is the event that no head comes up in all infinitely many tosses. Then we have $A_n \supset A_{n+1}$. So

$$P(A) = \lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} \frac{1}{2^n} = 0.$$