

概 率 论

Probability

肖磊, 2023年10月10日

已学知识点

● 第一章 事件与概率

▶ 随机现象与统计规律性

- ① 概率的频率解释依然是当今最通行的解释.
- ② 描述频率趋近于概率的大数定律总是概率论的第一大数定律.
- ③ 实际当中用频率作为概率的估计是十分自然的.

▶ 样本空间与事件

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω
$A \subset B$	A 是 B 的子集	A 发生则 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A - B$	差集	A 发生而 B 不发生
\bar{A}	余集	A 不发生

已学知识点

● 第一章 事件与概率

- ▶ 古典概型 (等可能概率模型): (1) 样本空间样本点有限; (2) 每个样本点等可能出现.
 - 计数方法: 排列组合.
 - 三个基本性质: 非负性、规范性、有限可加性.
- ▶ 几何概率: 以等可能性定义概率, 处理无限场合, 概率是几何体的测度之比.
 - 三个基本性质: 非负性、规范性、可列可加性.
- ▶ 概率空间: (Ω, \mathcal{F}, P)
 - 难点和要点: 事件域 \mathcal{F} 的选择, 太小不能满足需要, 太大难以定义概率.
 - 选择包含我们关注的所有事件的 σ 域, 保证事件对交、并、逆、差作可列次运算的封闭性.
 - 在这种 σ 域上, 能定义满足非负、规范和可列可加性的概率测度.

已学知识点

● 第二章 条件概率与统计独立性

▶ 条件概率的定义: $P(A|B) = \frac{P(AB)}{P(B)}$.

▶ 条件概率也是概率:

① $P(A|B) \geq 0$

② $P(\Omega|B) = 1$

③ $P\left(\bigcup_{i=1}^{\infty} A_i | B\right) = \sum_{i=1}^{\infty} P\left(A_i | B\right)$

A_i 两两互不相容

④ $P(\bar{A}|B) = 1 - P(A|B)$

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

已学知识点

● 第二章 条件概率与统计独立性

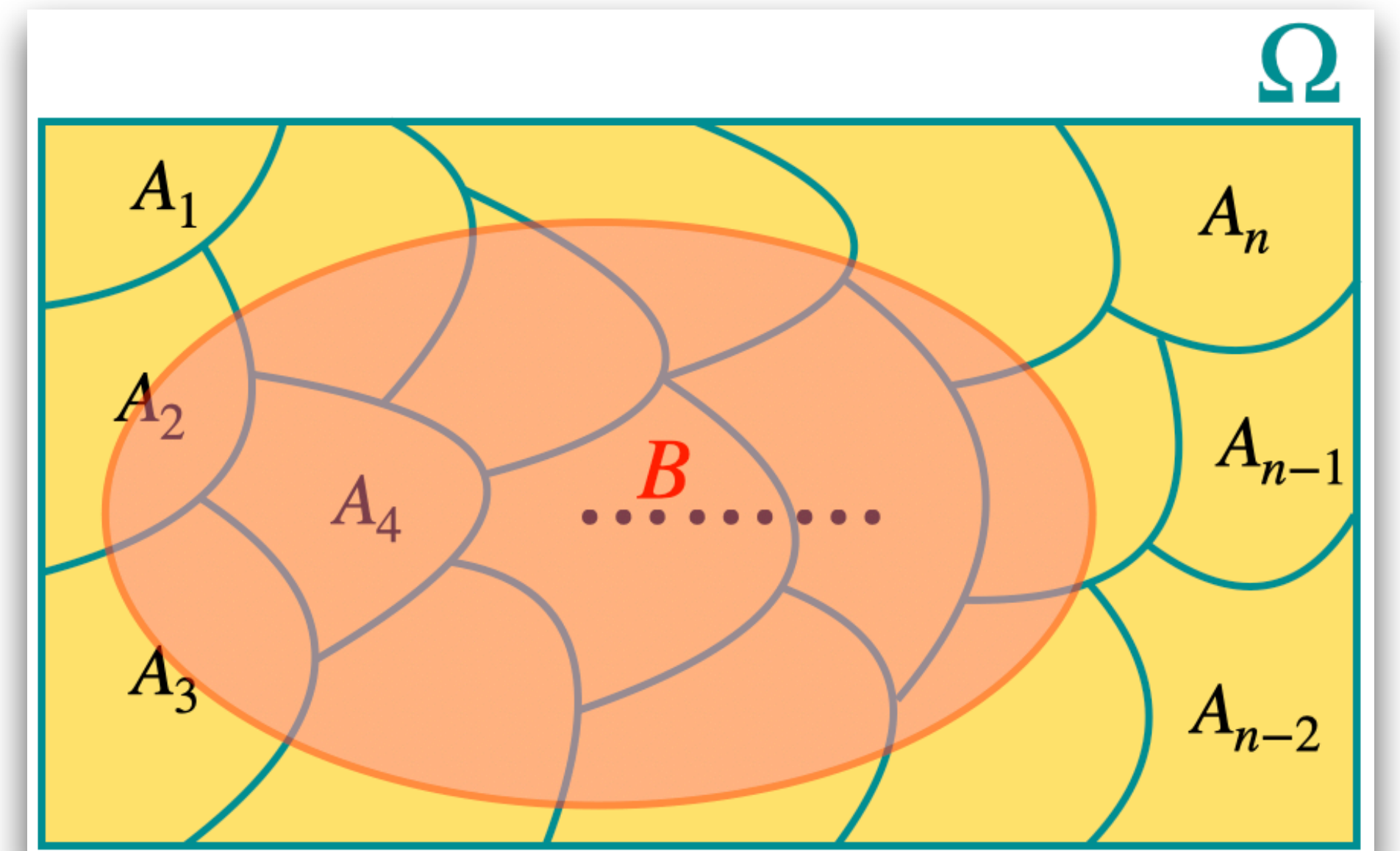
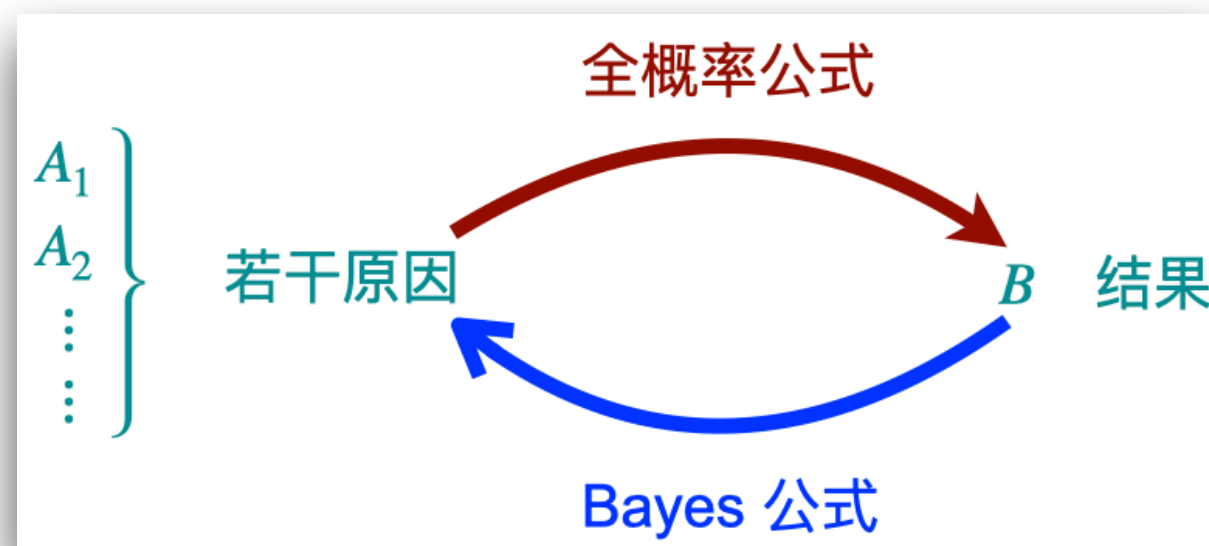
▶ 乘法公式: $P(AB) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$

$$P(A_1A_2 \cdots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1A_2) \cdots P(A_n|A_1A_2 \cdots A_{n-1})$$

乘法公式主要用于计算若干个事件同时发生的概率。

▶ 全概率公式: $P(B) = \sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i)$

▶ Bayes 公式: $P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i)}$



已学知识点

- 第二章 条件概率与统计独立性

- ▶ 两个事件独立: $P(AB) = P(A) \cdot P(B)$.

$\left. \begin{array}{l} \bar{A}, B \\ A, \bar{B} \\ \bar{A}, \bar{B} \end{array} \right\}$ 均相互独立

- ▶ 三个事件独立:
$$\left\{ \begin{array}{l} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{array} \right.$$

- ▶ A_1, A_2, \dots, A_n 相互独立, 则

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - [1 - P(A_1)] [1 - P(A_2)] \cdots [1 - P(A_n)]$$

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 试验相互独立：一个试验的结果对其它各试验的可能结果的概率都无影响.

- ▶ 试验独立性的定义：设 A_i 是试验 E_i 中的任一事件， $i = 1, 2, \dots, n$ ，若

$$P(A_1 A_2 \cdots A_n) = P(A_1) \cdot P(A_2) \cdot \cdots \cdot P(A_n)$$

则称试验 E_1, E_2, \dots, E_n 是相互独立的.

- ▶ 重复独立试验：研究在相同条件下重复进行的独立试验的数学模型.

已学知识点

- 第二章 条件概率与统计独立性

- ▶ Bernoulli 试验 E : 概率空间 (Ω, \mathcal{F}, P) , 其中

$$A \subset \Omega, \mathcal{F} = \{\Phi, A, \bar{A}, \Omega\}, P(A) = p, P(\bar{A}) = q, (p > 0, q > 0, p + q = 1)$$

- ▶ n 重 Bernoulli 试验 E_n : n 次独立重复的 Bernoulli 试验

- Bernoulli 分布: $n = 1$ 的 Bernoulli 试验.

$$P_1(k) = p^k q^{1-k}, k = 0, 1$$

样本点	H	T
概率	p	$1 - p$

- 二项分布: n 重 Bernoulli 试验中, 成功 (事件 A) 出现的次数

$$b(k; n, p) \triangleq P_n(k) = C_n^k p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, 2, \dots, n$$

已学知识点

- 第二章 条件概率与统计独立性

- ▶ Bernoulli 试验 E : 概率空间 (Ω, \mathcal{F}, P) , 其中

$$A \subset \Omega, \quad \mathcal{F} = \{\Phi, A, \bar{A}, \Omega\}, \quad P(A) = p, P(\bar{A}) = q, \quad (p > 0, q > 0, p + q = 1)$$

- ▶ n 重 Bernoulli 试验 E_n : n 次独立重复的 Bernoulli 试验

- 几何分布: n 重 Bernoulli 试验中, 首次成功出现在第 k 次

$$g(k; p) = q^{k-1}p = (1-p)^{k-1}p, \quad k = 1, 2, \dots$$

- Pascal 分布: n 重 Bernoulli 试验中, 第 r 次成功出现在第 k 次

$$f(k; r, p) = C_{k-1}^{r-1}p^r q^{k-r} = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, r+2, \dots$$

- 应用: 直线上的随机游动 (无限制、带有吸收壁)

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 二项分布推广到 n 个重复独立试验、且每次试验有若干可能结果的情形.

- 多项分布：每次试验的可能结果为 A_1, A_2, \dots, A_r ，且
$$\begin{cases} P(A_i) = p_i, & i = 1, 2, \dots, r \\ p_i \geq 0, & \text{且 } p_1 + p_2 + \dots + p_r = 1 \end{cases}$$

$$P(A_1 = k_1, A_2 = k_2, \dots, A_r = k_r) = \frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} p_2^{k_2} \cdot \dots \cdot p_r^{k_r}, \quad k_i \geq 0 \text{ 且 } k_1 + k_2 + \dots + k_r = n$$

- 应用：平面上的随机游动

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 二项分布： n 重 Bernoulli 试验中，事件 A 恰好发生 k 次的概率

$$b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n$$

- 概率计算：`dbinom(k, n, p)`

- 特点： k 增加时，概率 $b(k; n, p)$ 先随之增加直至达到最大值，随后单调减少。

当 $(n+1) \cdot p$ 不为整数时， $b(k; n, p)$ 在 $\lfloor (n+1) \cdot p \rfloor$ 达到最大值。

当 $(n+1) \cdot p = m$ 为整数时， $b(k; n, p)$ 在 $k = m - 1$ 和 $k = m$ 达到最大值。

一、二项分布

● **例：**(机票超售) 某航线历史资料表明：购票乘客有 5% 不来登机，问一架 200 座飞机可以售出多少张机票？

- ▶ 假设超售票 m 张，则共出售了 $200 + m$ 张机票.
- ▶ 每位乘客登机与否为 Bernoulli 试验，登机的概率 $p = 0.95$.
- ▶ 实际登机人数 k 人的概率服从二项分布 $b(k; 200 + m, 0.95)$.
- ▶ 出现拒登机的概率则为

$$\sum_{k=201}^{200+m} b(k; 200 + m, 0.95) = \sum_{k=201}^{200+m} \binom{200 + m}{k} \times 0.95^k \times 0.05^{200+m-k}$$

一、二项分布

- **例：**(机票超售) 某航线历史资料表明：购票乘客有 5% 不来登机，问一架 200 座飞机可以售出多少张机票？

```
f = function(m = 1, p = 0.95){  
  n = 200 + m  
  k = 201 : n  
  return(sum(dbinom(k, n, p)))  
}  
P = c()  
m = 20  
for (i in 1:m) {  
  P[i] = f(i)  
}  
x = data.frame(m = 1:m, Prob = P)  
round(x, digits = 6)
```

```
> round(x, digits = 6)  
   m   Prob  
1  1 0.000033  
2  2 0.000368  
3  3 0.002058  
4  4 0.007776  
5  5 0.022357  
6  6 0.052248  
7  7 0.103561  
8  8 0.179431  
9  9 0.278062  
10 10 0.392583  
11 11 0.512830  
12 12 0.628159  
13 13 0.730032  
14 14 0.813489  
15 15 0.877275  
16 16 0.922988  
17 17 0.953844  
18 18 0.973538  
19 19 0.985463  
20 20 0.992336
```

一、二项分布

● **例：**(人寿保险) 某年龄段参加保险者当中，假设一年中每个人死亡的概率为 0.005，有 10000 个这类人参加人寿保险，求未来一年中参加保险者里面

① 有 40 个人死亡的概率.

② 死亡人数不超过 70 人的概率.

▶ Bernoulli 概型， $n = 10000$ ， $p = 0.005$ 的二项分布.

① 有 40 个人死亡的概率.

$$b(40; 10000, 0.005) = \binom{10000}{40} \times 0.005^{40} \times 0.995^{9960} = 0.021434812$$

```
dbinom(40, 10000, 0.005)
```

```
> dbinom(40, 10000, 0.005)
[1] 0.021434812
```

一、二项分布

● **例：**(人寿保险) 某年龄段参加保险者当中，假设一年中每个人死亡的概率为 0.005，有 10000 个这类人参加人寿保险，求未来一年中参加保险者里面

① 有 40 个人死亡的概率.

② 死亡人数不超过 70 人的概率.

▶ Bernoulli 概型, $n = 10000$, $p = 0.005$ 的二项分布.

② 死亡人数不超过 70 人的概率.

$$\sum_{k=0}^{70} b(k; 10000, 0.005) = \sum_{k=0}^{70} \binom{10000}{k} \times 0.005^k \times 0.995^{10000-k} = 0.99709705$$

```
sum(dbinom(0:70, 10000, 0.005))
```

```
> sum(dbinom(0:70, 10000, 0.005))  
[1] 0.99709705
```

直接计算相当困难，
需要更有效的方法.

二、二项分布的 Poisson 逼近

Poisson 定理 在独立试验中，以 p_n 代表事件 A 在试验中出现的概率，它与试验的总次数 n 有关，如果 $np_n \rightarrow \lambda$ ，则当 $n \rightarrow \infty$ 时，有

$$b(k; n, p_n) \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\begin{aligned}
 b(k; n, p_n) &= \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\
 \lambda_n \triangleq np_n &\implies p_n = \frac{\lambda_n}{n} \\
 \text{已知 } \lim_{n \rightarrow \infty} \lambda_n &= \lambda \\
 \text{计算得 } \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^{n-k} &= e^{-\lambda} \\
 &= \frac{\lambda_n^k}{k!} \cdot \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\
 &= \frac{\lambda_n^k}{k!} \cdot \underbrace{\frac{n}{n} \frac{n-1}{n} \frac{n-2}{n} \cdots \frac{n-k+1}{n}}_k \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\
 &\xrightarrow{\quad} \frac{\lambda^k}{k!} e^{-\lambda} \quad (n \rightarrow \infty)
 \end{aligned}$$

二、二项分布的 Poisson 逼近

Poisson 定理 在独立试验中，以 p_n 代表事件 A 在试验中出现的概率，它与试验的总次数 n 有关，如果 $np_n \rightarrow \lambda$ ，则当 $n \rightarrow \infty$ 时，有

$$b\left(k; 10, \frac{3}{20}\right) \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}$$

$P\left(k; \frac{3}{2}\right)$

- 泊松定理表明：泊松分布是二项分布的极限分布。

应用中，当 p 很小 ($p \leq 0.1$) 时，我们有近似公式 $b(k; n, p) \approx \frac{(np)^k}{k!} e^{-np}$. } $n = 10, p = \frac{3}{20}$

n = 10

p = 3/20

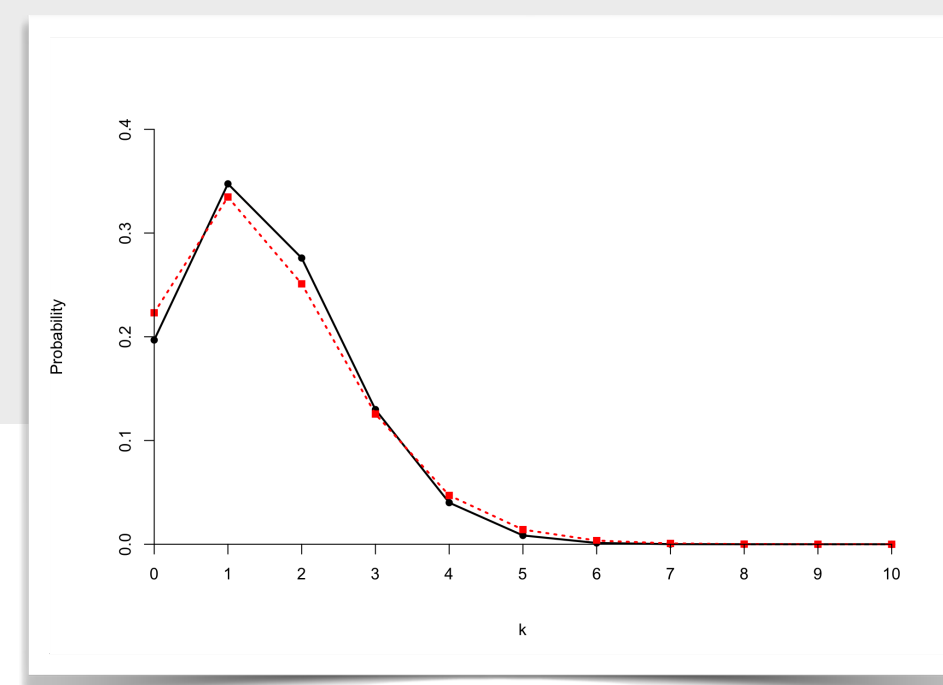
```
plot(0:10, dbinom(0:n, n, p), type = 'o', ylim = c(0, 0.4), pch = 16, xlab = 'k', ylab = 'Probability', lwd = 2, axes = FALSE)
```

```
lines(0:10, dpois(0:10, 3/2), col = 'red', lwd = 2, lty = 3)
```

```
points(0:10, dpois(0:10, 3/2), col = 'red', pch = 15)
```

```
axis(1, at = 0:10, pos = 0)
```

```
axis(2, pos = 0)
```



二、二项分布的 Poisson 逼近

Poisson 定理 在独立试验中，以 p_n 代表事件 A 在试验中出现的概率，它与试验的总次数 n 有关，如果 $np_n \rightarrow \lambda$ ，则当 $n \rightarrow \infty$ 时，有

$$b(k; n, p_n) \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}$$

- 近似实例：

$$\begin{cases} b(5; 500, 0.01) = 0.176351 \\ P(5; 500 \times 0.01) = 0.1754674 \end{cases}$$

$$\begin{cases} b(40; 10000, 0.005) = 0.02143481 \\ P(40; 10000 \times 0.005) = 0.02149963 \end{cases}$$

```
dbinom(5, 500, 0.01)
dpois(5, 5)
```

```
> dbinom(5, 500, 0.01)
[1] 0.176351
> dpois(5, 5)
[1] 0.1754674
```

```
> dbinom(40, 10000, 0.005)
[1] 0.02143481
> dpois(40, 50)
[1] 0.02149963
```

```
dbinom(40, 10000, 0.005)
dpois(40, 50)
```

三、Poisson (泊松) 分布

- 近几十年来, Poisson 分布日益显示其重要性, 成为概率论中最重要的几个分布之一. Poisson 分布在管理科学、运筹学以及自然科学的某些问题中都占有重要的地位.
 - ① 在大量重复试验中稀有事件出现的次数近似服从 Poisson 分布, 如意外事故, 罕见病, 大的自然灾害等.
 - ② 排队问题: 在一段时间内窗口等待服务的顾客数, 电话交换台的呼叫数, 网站访问数, 公共汽车站来到的乘客数等.
 - ③ 放射源衰变产生的粒子数.

三、Poisson (泊松) 分布

- 图形特点:

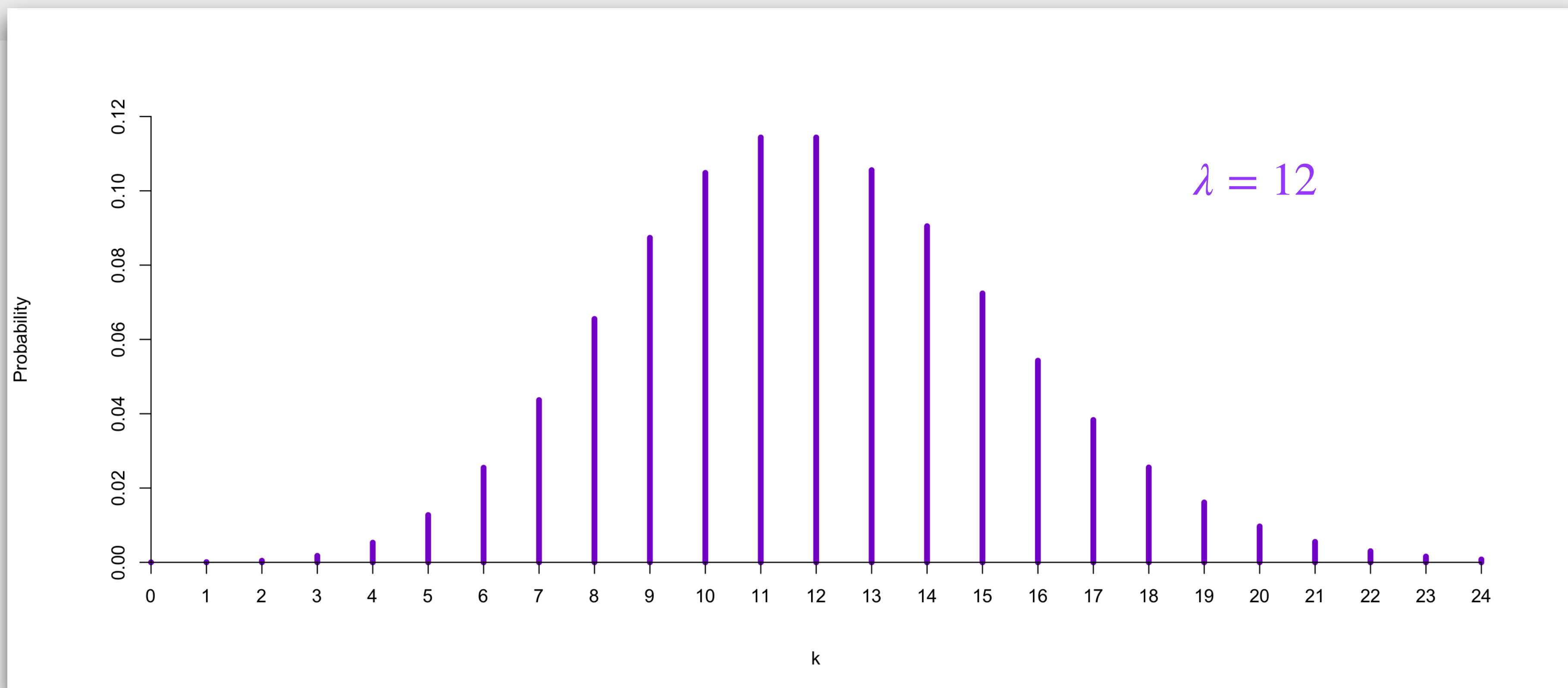
$$p(k; \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

x = 0 : 24

```
plot(x, dpois(x, 12), type = 'h', lwd = 5, axes = FALSE, xlab = 'k', ylab = 'Probability', ylim = c(0, 0.12), col = 'purple3')
```

```
axis(1, at = 0:24, pos = 0)
```

```
axis(2, pos = 0)
```



三、Poisson (泊松) 分布

- 例：**对上海市某公共汽车站的客流进行调查，统计了某天上午 10:30 至 11:47 左右每隔 20 秒钟来到的乘客批数 (每批可能有数人同时来到)，共得 230 个记录，分别计算了来到 0 批、1 批、2 批、3 批、4 批以及 4 批以上乘客的时间区间频数，结果如下

来到批数 i	0	1	2	3	≥ 4	总 共
频数 n_i	100	81	34	9	6	230
频率 $f_i = \frac{n_i}{n}$	0.43	0.35	0.15	0.04	0.03	
$p_i = \frac{\lambda^i}{i!}e^{-\lambda}$	0.42	0.36	0.16	0.05	0.01	

lambda = 0.87

round(dpois(0:4, lambda), digits = 2)

```

> lambda = 0.87
> round(dpois(0:4, lambda), digits = 2)
[1] 0.42 0.36 0.16 0.05 0.01
    
```

三、Poisson (泊松) 分布

- 例：**放射性物质放射出的 α 质点数是服从 Poisson 分布的著名例子. 1910 年 Rutherford 等人的实验揭露了这个事实. 在这个实验中，观察了长为 7.5 秒的时间间隔里到达某指定区域的质点数，共观察了 $N = 2608$ 次. 对照数据如下

```

k = 0:10
N_k = c(57, 203, 383, 525, 532, 408, 273, 139, 45, 27, 16)
Expected = dpois(k, 3.87) * 2608
Expected[11] = (1 - sum(dpois(0:9, 3.87))) * 2608
x = data.frame(k = k, N_k = N_k, Expect = Expected)
round(x, digit = 3)
    
```

```

> round(x, digit = 3)
   k N_k  Expect
1  0  57  54.399
2  1 203 210.523
3  2 383 407.361
4  3 525 525.496
5  4 532 508.418
6  5 408 393.515
7  6 273 253.817
8  7 139 140.325
9  8  45  67.882
10 9  27  29.189
11 10+ 16  17.075
    
```

- ▶ 理论值与实验值很近似.

三、Poisson (泊松) 分布

- 产生 Poisson 分布的机制分析

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

$$x = \frac{t}{2}, y = \frac{t}{2}$$

$$f(t) = \left[f\left(\frac{t}{2}\right) \right]^2 \geq 0$$

非负

Cauchy 引理 若 $f(x)$ 是连续函数 (或单调函数), 且对一切

x, y (或一切 $x \geq 0, y \geq 0$) 成立

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

则

$$f(x) = a^x$$

其中 $a \geq 0$ 是某一常数.

$$f(x) \cdot f(y) = f(x + y) \xrightarrow{y=x} f(2x) = [f(x)]^2$$

$$\xrightarrow{y=2x} f(3x) = f(x) \cdot f(2x) = [f(x)]^3 \implies f(nx) = [f(x)]^n \quad x = \frac{1}{n}$$

$$a^{\frac{m}{n}} = \left[f\left(\frac{1}{n}\right) \right]^m = \underbrace{f\left(\frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}\right)}_m = f\left(\frac{m}{n}\right) \longleftarrow f\left(\frac{1}{n}\right) = a^{\frac{1}{n}} \longleftarrow 0 \leq a \triangleq f(1) = \left[f\left(\frac{1}{n}\right) \right]^n$$

► 说明 $f(x) = a^x$ 对一切有理数成立. 再利用连续性或单调性可以证明对无理数也成立.

三、Poisson (泊松) 分布

概率规律不随时间的推移而改变

互不相交的时间区间内
过程进行的相互独立性

● **Poisson 过程**: 考虑来到某交换装置的电话呼叫数, 假定它具有下面三个性质

- ① **平稳性**: 在 $[t_0, t_0 + t)$ 中来的呼叫数只与时间间隔长度 t 有关, 而与时间的起点 t_0 无关.
- ② **独立增量性 (无后效性)**: 在 $[t_0, t_0 + t)$ 中来到 k 次呼叫与时刻 t_0 以前发生的事件独立.
- ③ **普通性**: 在充分小的时间间隔内, 最多来到一个呼叫.

$N(t) \triangleq$ 长度为 t 的时间区间中来的呼叫数

$P_k(t) \triangleq P(N(t) = k)$

$$\} \implies \sum_{k=0}^{\infty} P_k(t) = 1$$

在同一瞬间不可能来到
两个或两个以上呼叫

$$\varphi(t) \triangleq \sum_{k=2}^{\infty} P_k(t) = 1 - P_0(t) - P_1(t) = o(t) \implies \lim_{t \rightarrow 0} \frac{\varphi(t)}{t} = 0$$

长度为 t 的时间区间来到 2 次及以上呼叫的概率

三、Poisson (泊松) 分布

- 下面求 $P_k(t)$: 对于 $\Delta t > 0$, 考虑 $[0, t + \Delta t)$ 中发生 k 次事件的概率 $P_k(t + \Delta t)$.



独立增量性 }
全概率公式 }

$$\Rightarrow P_k(t + \Delta t) = P_k(t)P_0(\Delta t) + P_{k-1}(t)P_1(\Delta t) + P_{k-2}(t)P_2(\Delta t) + \dots + P_0(t)P_k(\Delta t), \quad k \geq 0$$

$k = 0$

$$\rightarrow P_0(t + \Delta t) = P_0(t)P_0(\Delta t)$$

- $P_0(t)$ 是长为 t 的时间区间中无呼叫的概率, 随着时间区间长度 t 的增大, 这一概率会减小, 所以 $P_0(t)$ 关于 t 单调下降. 由 Cauchy 引理

$$P_0(t) = a^t$$

$a \geq 0$ 为常数

$$P_0(t) = a^t \text{ 为概率 } \Rightarrow 0 \leq a \leq 1$$

Cauchy 引理 若 $f(x)$ 是连续函数 (或单调函数), 且对一切 x, y (或一切 $x \geq 0, y \geq 0$) 成立

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

则

$$f(x) = a^x$$

其中 $a \geq 0$ 是某一常数.

三、Poisson (泊松) 分布

- 下面求 $P_k(t)$: 对于 $\Delta t > 0$, 考虑 $[0, t + \Delta t)$ 中发生 k 次事件的概率 $P_k(t + \Delta t)$.
 - ▶ $a = 0$ 时, $P_0(t) = 0$, 说明无论时间间隔 t 多么小, 无呼叫的概率为 0, 即都会有呼叫, 此时, 有限时间间隔区间内会有无限个呼叫, 不考虑!
 - ▶ $a = 1$ 时, $P_0(t) = 1$, 说明永远无呼叫, 不考虑!

} $\implies 0 < a < 1$

$$\lambda \triangleq -\ln a > 0$$

$$\implies a = e^{-\lambda}, \quad (\lambda > 0)$$

$a \geq 0$ 为常数

$$P_0(t) = a^t$$

$P_0(t) = a^t$ 为概率 $\implies 0 \leq a \leq 1$

$$\implies P_0(t) = e^{-\lambda t}$$

三、Poisson (泊松) 分布

- 下面求 $P_k(t)$: 对于 $\Delta t > 0$, 考虑 $[0, t + \Delta t)$ 中发生 k 次事件的概率 $P_k(t + \Delta t)$.

▶ 当 $\Delta t \rightarrow 0$ 时,

在 0 点多项式展开

$$P_0(\Delta t) = e^{-\lambda\Delta t} = 1 - \lambda\Delta t + o(\Delta t) \implies 1 - P_0(\Delta t) = \lambda\Delta t + o(\Delta t)$$

$$P_1(\Delta t) = 1 - P_0(\Delta t) - P[N(\Delta t) \geq 2]$$

由普通性知 $P[N(\Delta t) \geq 2] = o(\Delta t)$

$$P_{k-l}(t) \leq 1 \quad = \lambda\Delta t + o(\Delta t)$$

$$\sum_{l=2}^{\infty} P_{k-l}(t) \cdot P_l(\Delta t) \leq \sum_{l=2}^{\infty} P_l(\Delta t) = P[N(\Delta t) \geq 2] = o(\Delta t)$$

$$P_k(t + \Delta t) = P_k(t)P_0(\Delta t) + P_{k-1}(t)P_1(\Delta t) + P_{k-2}(t)P_2(\Delta t) + \cdots + P_0(t)P_k(\Delta t), \quad k \geq 0$$

$$\implies P_k(t + \Delta t) = P_k(t) \cdot (1 - \lambda\Delta t) + P_{k-1}(t) \cdot \lambda\Delta t + o(\Delta t), \quad (k \geq 1)$$

$$\implies P_0(t) = e^{-\lambda t}$$

$$= P_k(t) - P_k(t) \cdot \lambda\Delta t + P_{k-1}(t) \cdot \lambda\Delta t + o(\Delta t)$$

三、Poisson (泊松) 分布

- 下面求 $P_k(t)$: 对于 $\Delta t > 0$, 考虑 $[0, t + \Delta t)$ 中发生 k 次事件的概率 $P_k(t + \Delta t)$.

▶ 于是

$$\frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = \lambda [P_{k-1}(t) - P_k(t)] + o(1), \quad (k \geq 1)$$

两侧同乘 $e^{\lambda t}$

$$\Delta t \rightarrow 0 \implies P'_k(t) = \lambda [P_{k-1}(t) - P_k(t)], \quad (k \geq 1)$$

$$\implies e^{\lambda t} P'_k(t) = \lambda e^{\lambda t} P_{k-1}(t) - \lambda e^{\lambda t} P_k(t), \quad (k \geq 1) \implies [e^{\lambda t} P_k(t)]' = \lambda e^{\lambda t} P_{k-1}(t), \quad (k \geq 1)$$

$$\xrightarrow{k=1} [e^{\lambda t} P_1(t)]' = \lambda e^{\lambda t} P_0(t) \quad \xrightarrow{P_0(t)=e^{-\lambda t}} [e^{\lambda t} P_1(t)]' = \lambda \implies P_1(t) = \lambda t e^{-\lambda t}$$

$$\implies P_k(t + \Delta t) = P_k(t) \cdot (1 - \lambda \Delta t) + P_{k-1}(t) \cdot \lambda \Delta t + o(\Delta t), \quad (k \geq 1)$$

$$= P_k(t) - P_k(t) \cdot \lambda \Delta t + P_{k-1}(t) \cdot \lambda \Delta t + o(\Delta t)$$

三、Poisson (泊松) 分布

- 下面求 $P_k(t)$: 对于 $\Delta t > 0$, 考虑 $[0, t + \Delta t)$ 中发生 k 次事件的概率 $P_k(t + \Delta t)$.

► 于是

$$\frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = \lambda [P_{k-1}(t) - P_k(t)] + o(1), \quad (k \geq 1)$$

两侧同乘 $e^{\lambda t}$

$$\Delta t \rightarrow 0 \implies P'_k(t) = \lambda [P_{k-1}(t) - P_k(t)], \quad (k \geq 1)$$

$$\implies e^{\lambda t} P'_k(t) = \lambda e^{\lambda t} P_{k-1}(t) - \lambda e^{\lambda t} P_k(t), \quad (k \geq 1) \implies [e^{\lambda t} P_k(t)]' = \lambda e^{\lambda t} P_{k-1}(t), \quad (k \geq 1)$$

$$\xRightarrow{k=1} [e^{\lambda t} P_1(t)]' = \lambda e^{\lambda t} P_0(t) \quad \begin{matrix} P_0(t) = e^{-\lambda t} \\ \implies \end{matrix} [e^{\lambda t} P_1(t)]' = \lambda \implies P_1(t) = \lambda t e^{-\lambda t}$$

$$\xRightarrow{k=2} [e^{\lambda t} P_2(t)]' = \lambda e^{\lambda t} P_1(t) \quad \begin{matrix} P_1(t) = \lambda t e^{-\lambda t} \\ \implies \end{matrix} [e^{\lambda t} P_2(t)]' = \lambda^2 t \implies P_2(t) = \frac{(\lambda t)^2}{2!} e^{-\lambda t}$$

$$\xRightarrow{k=3} [e^{\lambda t} P_3(t)]' = \lambda e^{\lambda t} P_2(t) \quad \begin{matrix} P_2(t) = \frac{(\lambda t)^2}{2!} e^{-\lambda t} \\ \implies \end{matrix} [e^{\lambda t} P_3(t)]' = \frac{\lambda^3 t^2}{2!} \implies P_3(t) = \frac{(\lambda t)^3}{3!} e^{-\lambda t}$$

三、Poisson (泊松) 分布

- 下面求 $P_k(t)$: 对于 $\Delta t > 0$, 考虑 $[0, t + \Delta t)$ 中发生 k 次事件的概率 $P_k(t + \Delta t)$.

▶ 于是

$$\frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = \lambda [P_{k-1}(t) - P_k(t)] + o(1), \quad (k \geq 1)$$

$$\Rightarrow P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \dots$$

参数为 λt 的 Poisson 分布

三、Poisson (泊松) 分布

- 例：假设服务器在长度 t 分钟的时间内被攻击的次数近似服从 $P(2t)$ ，问 3 分钟内至少被攻击一次与 5 分钟内至少被攻击两次，哪一个事件更有可能出现？

- ▶ 3 分钟内被攻击次数服从参数为 6 的 Poisson 分布，故 3 分钟内至少被攻击一次的概率为

$$1 - p(0; 6) \approx 0.9975212$$

```
1 - dpois(0, 6)
```

```
> 1 - dpois(0, 6)  
[1] 0.9975212
```

- ▶ 5 分钟内被攻击次数服从参数为 10 的 Poisson 分布，故 5 分钟内至少被攻击两次的概率为

$$1 - p(0; 10) - p(1; 10) \approx 0.9995006$$

```
1 - dpois(0, 10) - dpois(1, 10)
```

```
> 1 - dpois(0, 10) - dpois(1, 10)  
[1] 0.9995006
```

- ▶ 5 分钟内至少被攻击两次的事件更有可能会出现。

第三章 随机变量与分布函数

3.1 随机变量及其分布

3.2 随机向量, 随机向量的独立性

3.3 随机变量的函数及其分布

3.1 随机变量及其分布

- 一、随机变量及分布函数的定义
- 二、分布函数的性质
- 三、离散型随机变量
- 四、连续型随机变量

一、随机变量的定义

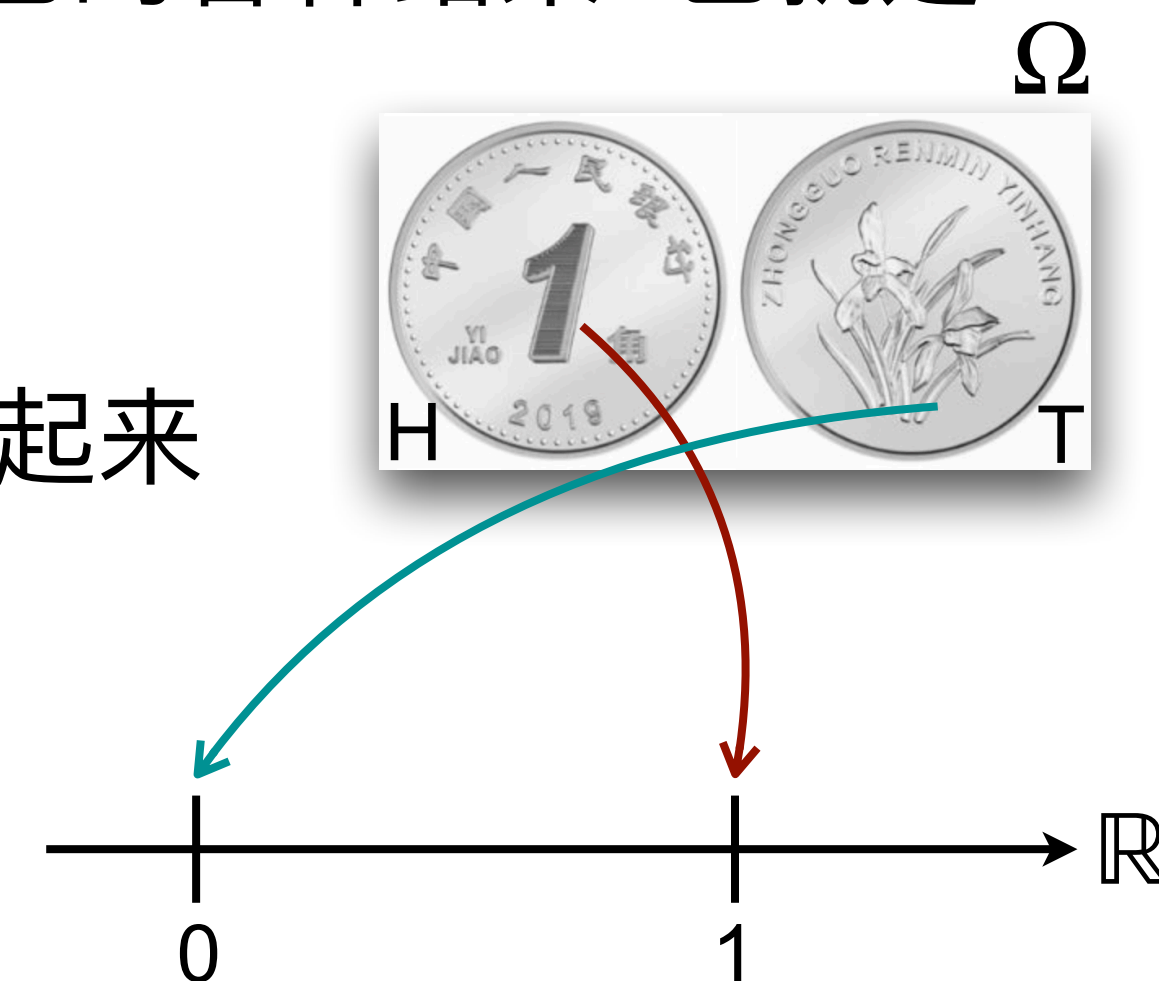
- 随机试验中，有些试验结果本身与数值有关 (本身就是一个数).
 - ▶ 掷一枚骰子，向上一面出现的点数.
 - ▶ 每天进出 C5 - 302 教室的人数.
 - ▶ 昆虫的产卵数.
 - ▶ 克拉玛依每年七月份的最高气温.

一、随机变量的定义

- 有些试验中，试验结果看起来与数值无关，但可以引进一个变量来表示它的各种结果. 也就是说，把试验结果数值化.

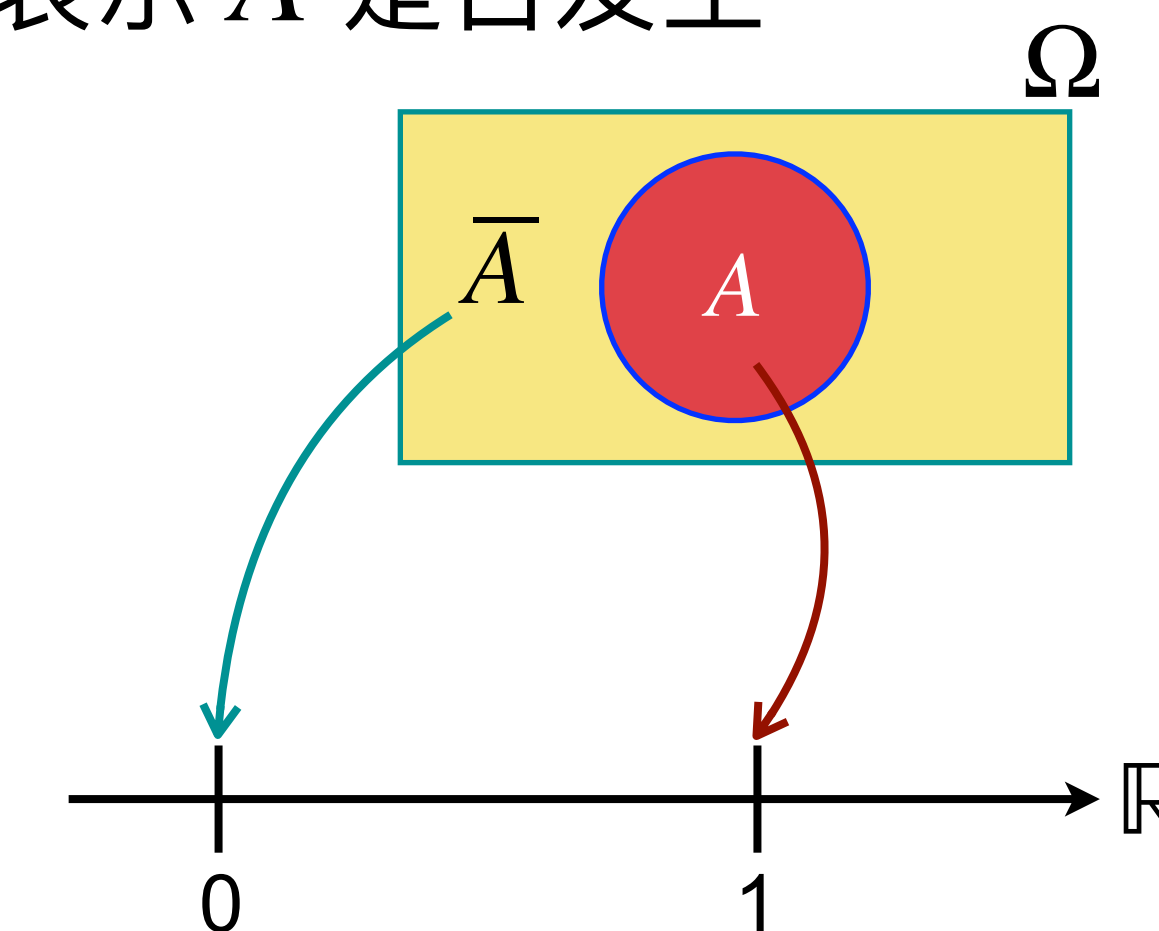
- ▶ 例：抛一枚硬币的试验中，可以通过如下方法将试验结果与数值联系起来

$$X = \begin{cases} 1, & \text{出现正面 (H) 向上} \\ 0, & \text{出现反面 (T) 向上} \end{cases}$$



- ▶ 例：一般，对任意随机事件 A ，可以引进一个函数 ($\Omega \rightarrow \{0, 1\}$) 来表示 A 是否发生

$$1_A = \begin{cases} 1, & \text{如果 } A \text{ 发生} \\ 0, & \text{如果 } A \text{ 不发生} \end{cases}$$



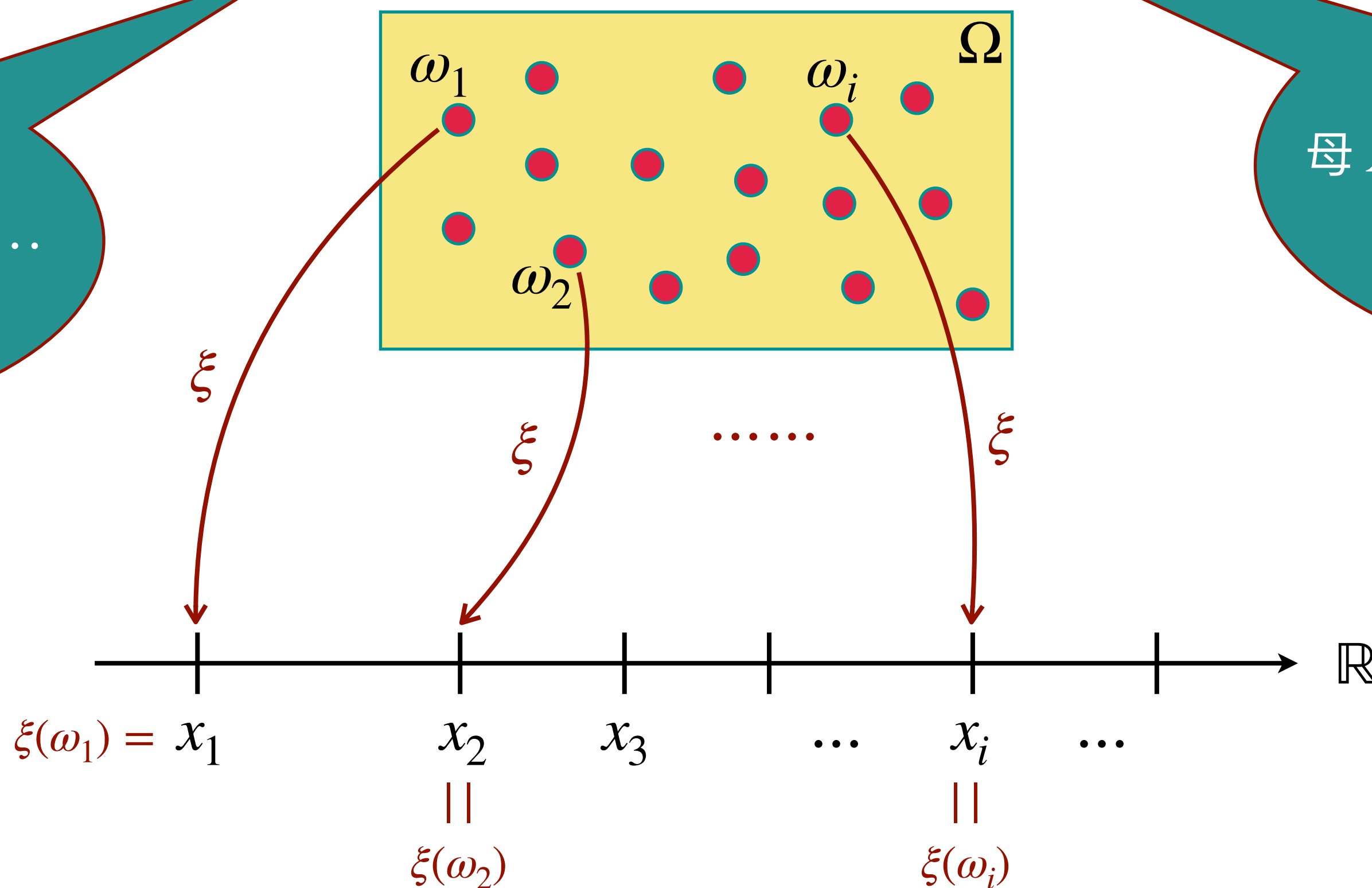
- ▶ 这种对应关系在数学上理解为定义了一种实值单值函数.

一、随机变量的定义

- 随机变量 (random variable, r.v.) 的概念

若对于随机试验 E 的每一个可能结果 $\omega \in \Omega$ ，都有唯一的一个实数值 $\xi(\omega)$ 相对应，则称实值函数 $\xi(\omega)$ 为随机变量，简记为 ξ 。

随机变量所取的值一般采用小写字母 x, y, z, w, v, \dots 等来表示



随机变量通常用大写字母 X, Y, Z, W, V, \dots 或希腊字母 ξ, η, ζ, \dots 等表示

一、随机变量的定义

- 如何理解一个随机变量 ξ ?

- ▶ ξ 的取值情况:

- ① 随机变量 ξ 是定义在样本空间上的实值函数.
- ② 它的取值与试验结果形成对应.
- ③ 随着实验结果的不同而取不同的值.
- ④ 在试验之前只知道 ξ 可能取值的范围, 而不能预先肯定它将取哪个值.

一、随机变量的定义

- 如何理解一个随机变量 ξ ?

- ▶ ξ 取值的概率分布情况:

- ① 由于试验结果的出现具有一定的概率, 所以随机变量取每个值和每个确定范围内的值也具有一定的概率.

- ② 随机变量的取值既具有可变性, 也有随机性. 这种双重性正是随机变量与普通变量 (函数) 的本质区别.

一、随机变量的定义

- 引入随机变量的意义?

研究**随机事件**的概率规律 $\xrightarrow{\text{通过将随机事件数值化转为}}$ 研究**随机变量取值**的概率规律

以函数为工具

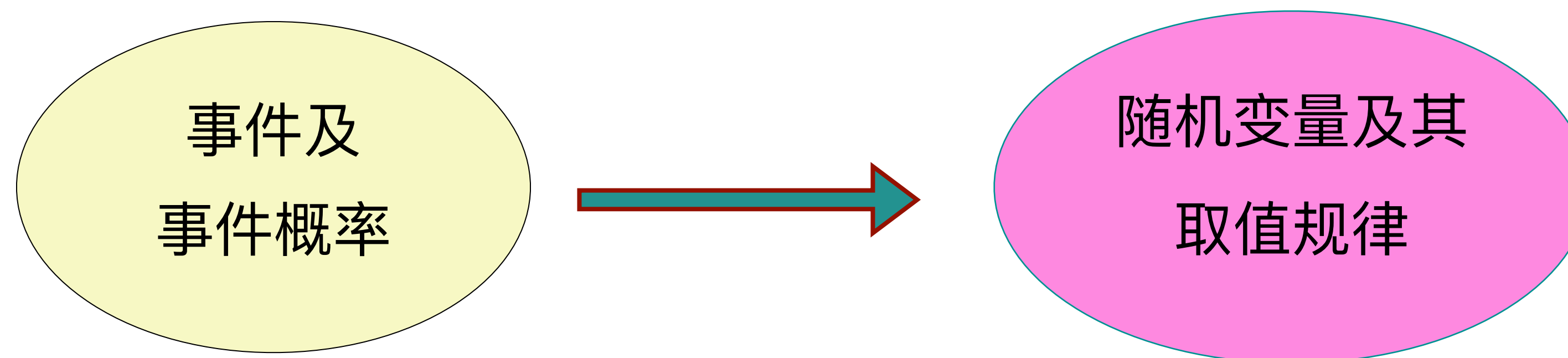
使概率可转化为我们所熟知的函数形式

分析工具有了用武之地

一、随机变量的定义

- 引入随机变量的意义?

随机变量概念的产生是概率论发展史上的重大事件. 引入随机变量后, 随机试验中的任一随机事件就可以通过随机变量的取值关系式表达出来, 对随机现象统计规律的研究, 就由对事件及事件概率的研究扩大为对随机变量及其取值规律的研究.



一、随机变量的定义

● 例：① 生化检验结果分阳性和阴性 $X = \begin{cases} 1, & \text{检验结果为阳性} \\ 0, & \text{检验结果为阴性} \end{cases}$

② 抛一枚硬币，结果分为正面或反面 $X = \begin{cases} 1, & \text{硬币正面向上} \\ 0, & \text{硬币反面向上} \end{cases}$

③ 抛掷一枚均匀的骰子，观察向上一面出现的点数



$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$\xi(\omega) = \omega \quad \downarrow \quad \text{恒等变换}$$

$$\xi(1) = 1, \xi(2) = 2, \xi(3) = 3, \xi(4) = 4, \xi(5) = 5, \xi(6) = 6$$

$$\Rightarrow P(\xi = i) = \frac{1}{6}, \quad i = 1, 2, 3, 4, 5, 6$$

一、随机变量的定义

- 定义**: 设 $\xi(\omega)$ 是定义于概率空间 (Ω, \mathcal{F}, P) 上的单值实函数, 如果对于直线上任一 Borel 点集 B , 有

$$\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$$

则称 $\xi(\omega)$ 为**随机变量**, 而 $P(\xi(\omega) \in B)$ 称为随机变量 $\xi(\omega)$ 的**概率分布**.

