概 率 论

Probability

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：$\left(\Omega, \mathscr{F}, P\right)$
    - 难点和要点：事件域 $\mathscr{F}$ 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：$P(A\,|\,B)=\dfrac{P(AB)}{P(B)}.$

    - 乘法公式：$P\left(A_1A_2\cdots A_n\right)=P\left(A_1\right)\cdot P\left(A_2\middle|\,A_1\right)\cdot P\left(A_3\middle|\,A_1A_2\right)\cdot\cdots\cdot P\left(A_n\middle|\,A_1A_2\cdots A_{n-1}\right)$
    - 全概率公式：$P(B)=\sum_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\,\middle|A_i\right)$
    - Bayes 公式：$P\left(A_k\,\middle|B\right)=\dfrac{P\left(A_k\right)\cdot P\left(B\,\middle|A_k\right)}{\sum_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\,\middle|A_i\right)}$

$$\left.\begin{array}{l} A_i\cap A_j=\Phi\quad(i\neq j)\\ \bigcup_{i=1}^{\infty}A_i=\Omega \end{array}\right\}$$

已学知识点

- 第二章 条件概率与统计独立性
  - 独立性：两个事件独立 $P(AB)=P(A)\cdot P(B)$.

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

  - 三个事件独立 $\begin{cases} P(AB)=P(A)\cdot P(B)\\ P(AC)=P(A)\cdot P(C)\\ P(BC)=P(B)\cdot P(C)\\ P(ABC)=P(A)\cdot P(B)\cdot P(C)\end{cases}$
  - $A_1, A_2, \ldots, A_n$ 相互独立，则

$$P\left(A_1\cup A_2\cup\cdots\cup A_n\right)=1-\left[1-P\left(A_1\right)\right]\left[1-P\left(A_2\right)\right]\cdot\cdots\cdot\left[1-P\left(A_n\right)\right]$$

  - 试验独立：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.

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 $\left(\Omega, \mathscr{F}, P\right)$，其中

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

    - n 重 Bernoulli 试验 E_n：n 次独立重复的 Bernoulli 试验
    - Bernoulli 分布：$P_1(k) = p^k q^{1-k}$ ， $k = 0, 1$
    - 二项分布：$b\left(k;\ n, p\right) \triangleq P_n(k) = C_n^k p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$ ， $k = 0, 1, 2, \ \ldots, n$
    - 几何分布：$g(k;\ p) = q^{k-1} p = (1-p)^{k-1} p$ ， $k = 1, 2, \ \ldots$
    - Pascal 分布：$f(k;\ r, p) = C_{k-1}^{r-1} p^r q^{k-r} = \binom{k-1}{r-1} p^r (1-p)^{k-r}$ ， $k = r, r+1, r+2, \ \ldots$
    - 动态模型：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

  - Bernoulli 试验 E：概率空间 $(\Omega, \mathscr{F}, P)$，其中

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

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

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

    - 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

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

    - 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 $\left[(n+1)\cdot p\right]$ 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.

已学知识点

- 第二章 条件概率与统计独立性
  - Poisson 分布：$p(k;\ \lambda) = \dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$, $k = 0, 1, 2, \ldots$
    - 概率计算：$\mathrm{dpois}\left(k, \lambda\right)$
    - Poisson 过程：

      ① 平稳性：在 $\left[t_0\ ,\ t_0 + t\right)$ 中发生的次数只与时间间隔长度 t 有关，而与起点 t_0 无关.

      ② 独立增量性 (无后效性)：在 $\left[t_0\ ,\ t_0 + t\right)$ 中发生 k 次与时刻 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，最多发生一次.

第三章 随机变量与分布函数

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

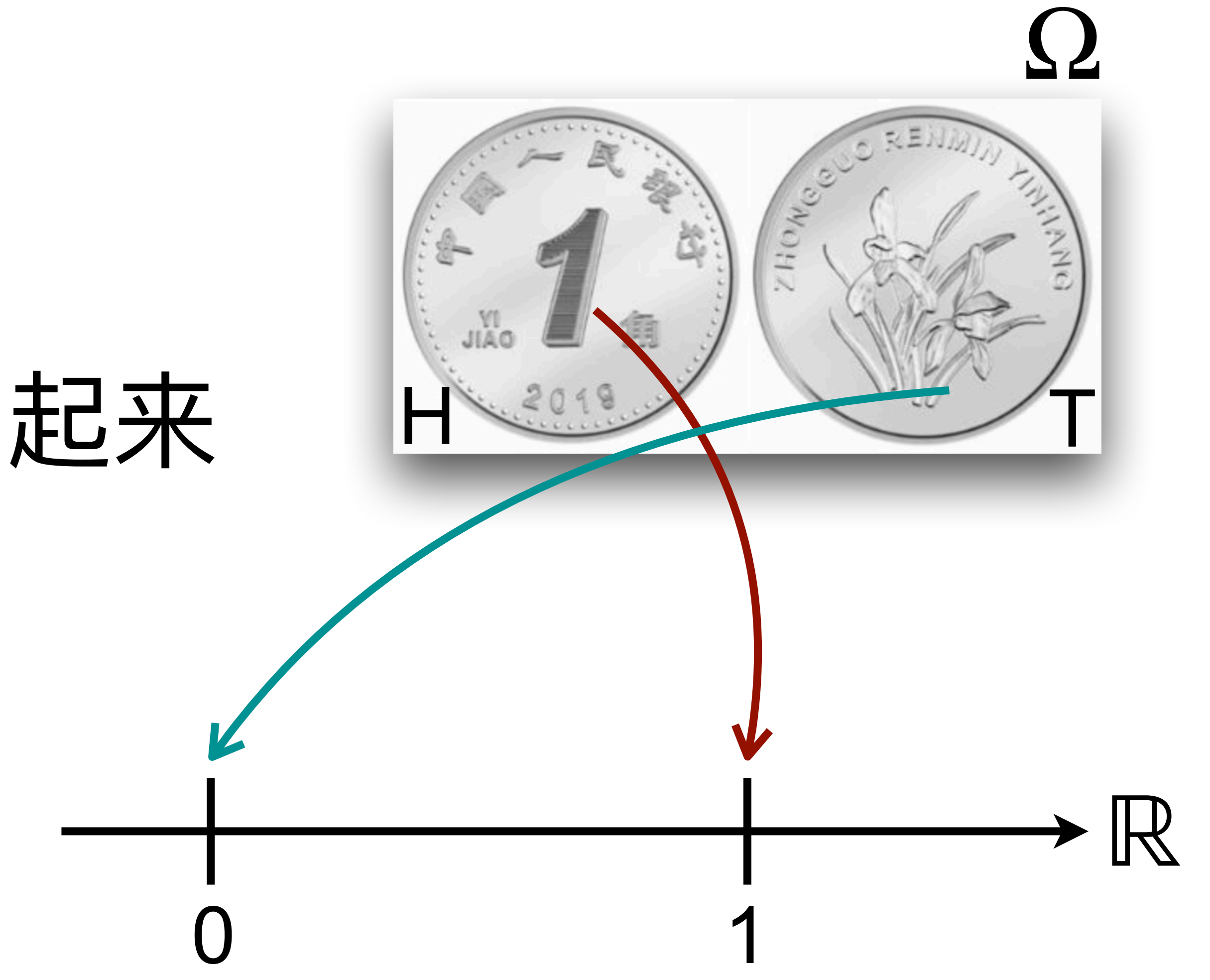
四、连续型随机变量

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 - 402 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

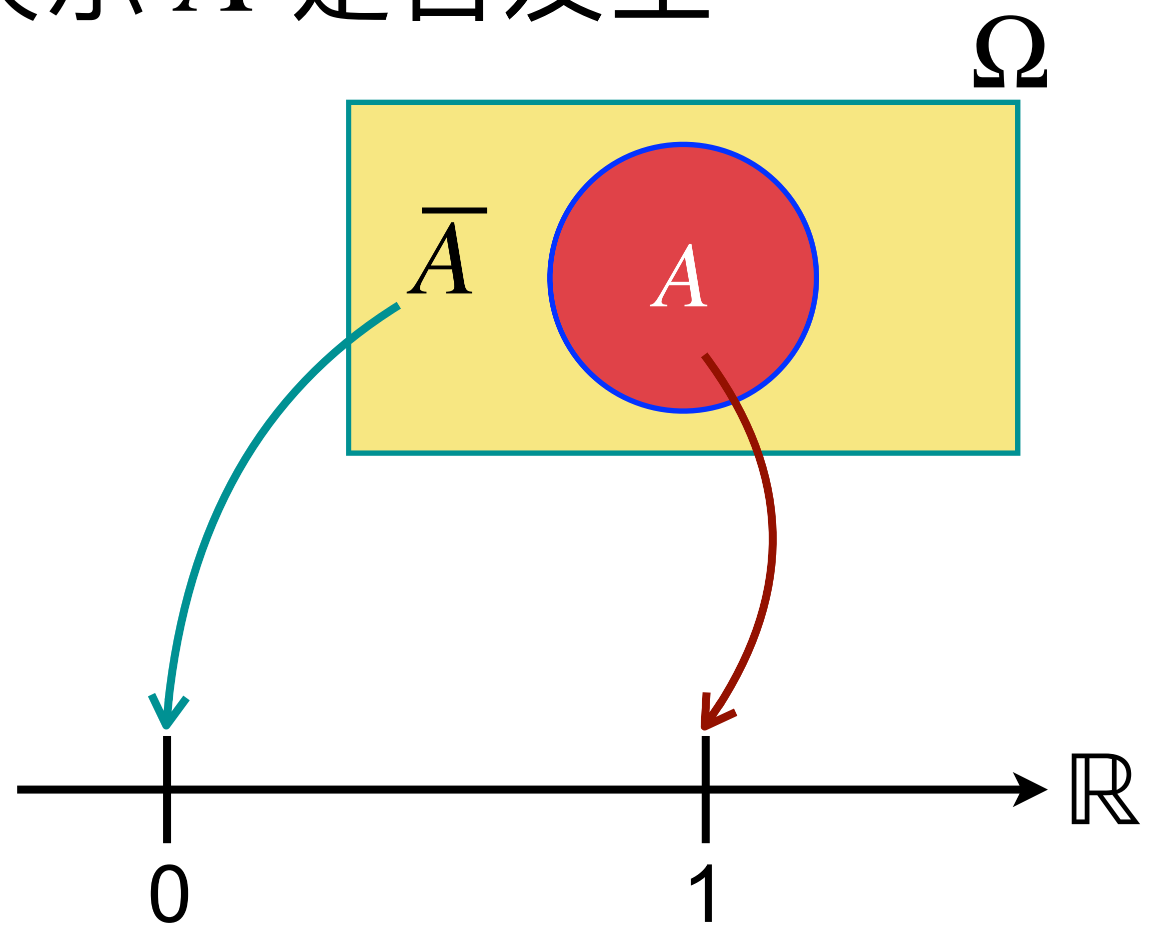
- 有些试验中，试验结果看起来与数值无关，但可以引进一个变量来表示它的各种结果. 也就是说，把试验结果数值化.



▶ 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 $\big(\Omega \to \{0\,,1\}\big)$ 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.

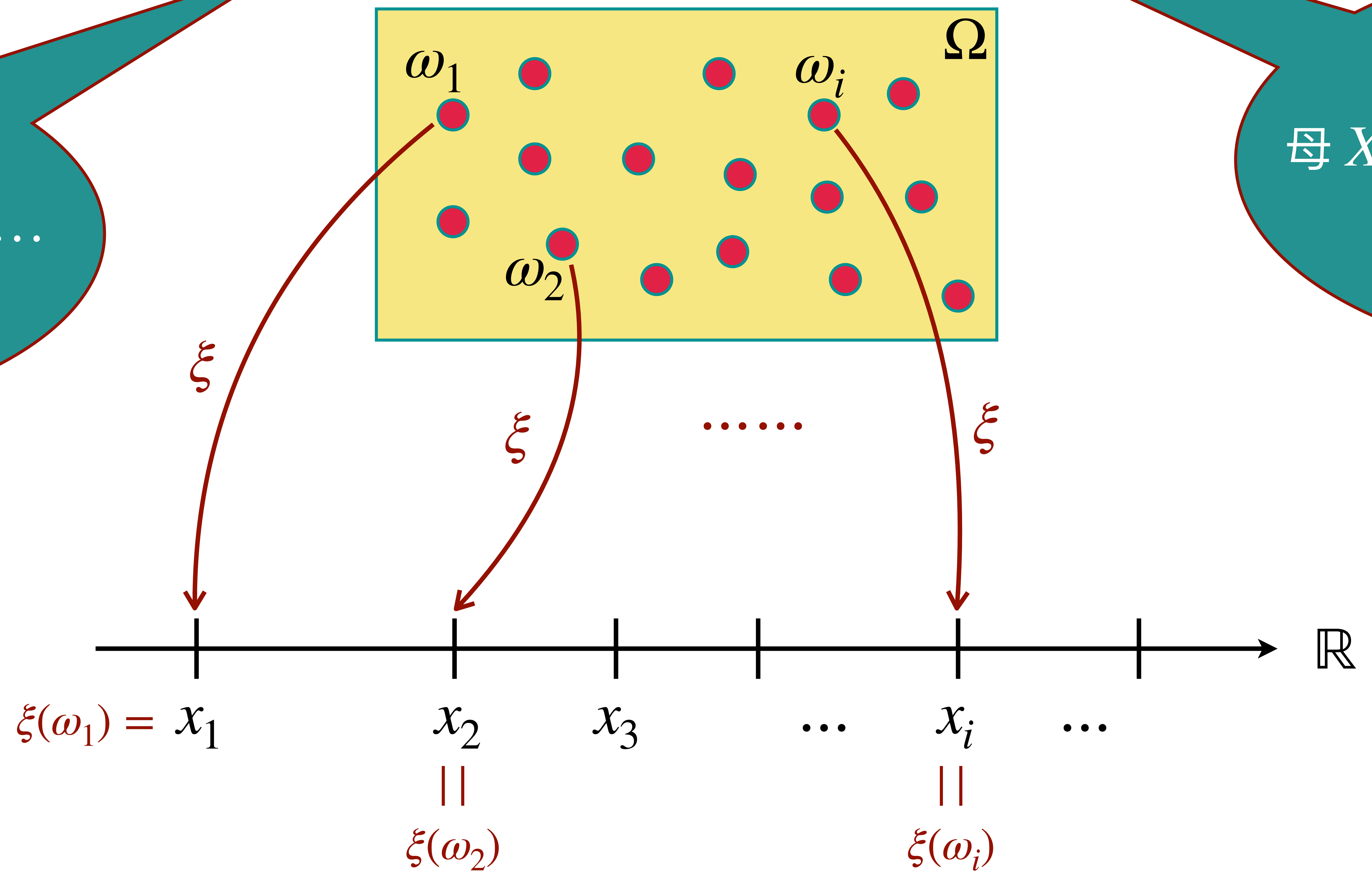
一、随机变量的定义

- 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 $x, y, z, w, v, \ldots$ 等来表示

随机变量通常用大写字母 $X, Y, Z, W, V, \ldots$ 或希腊字母 $\xi, \eta, \zeta, \ldots$ 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

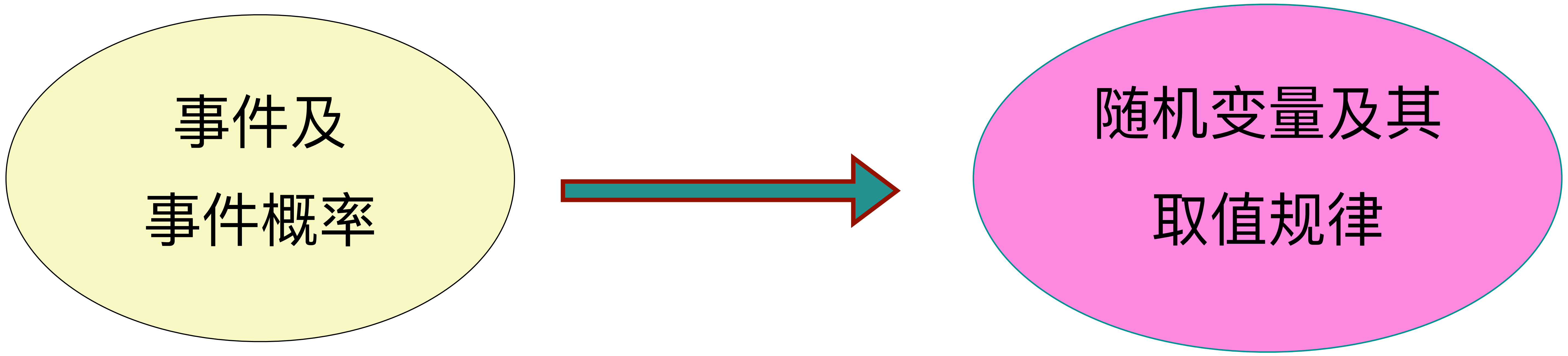
使概率可转化为我们所熟知的函数形式

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 

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

$\xi(\omega)=\omega$ 恒等变换

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

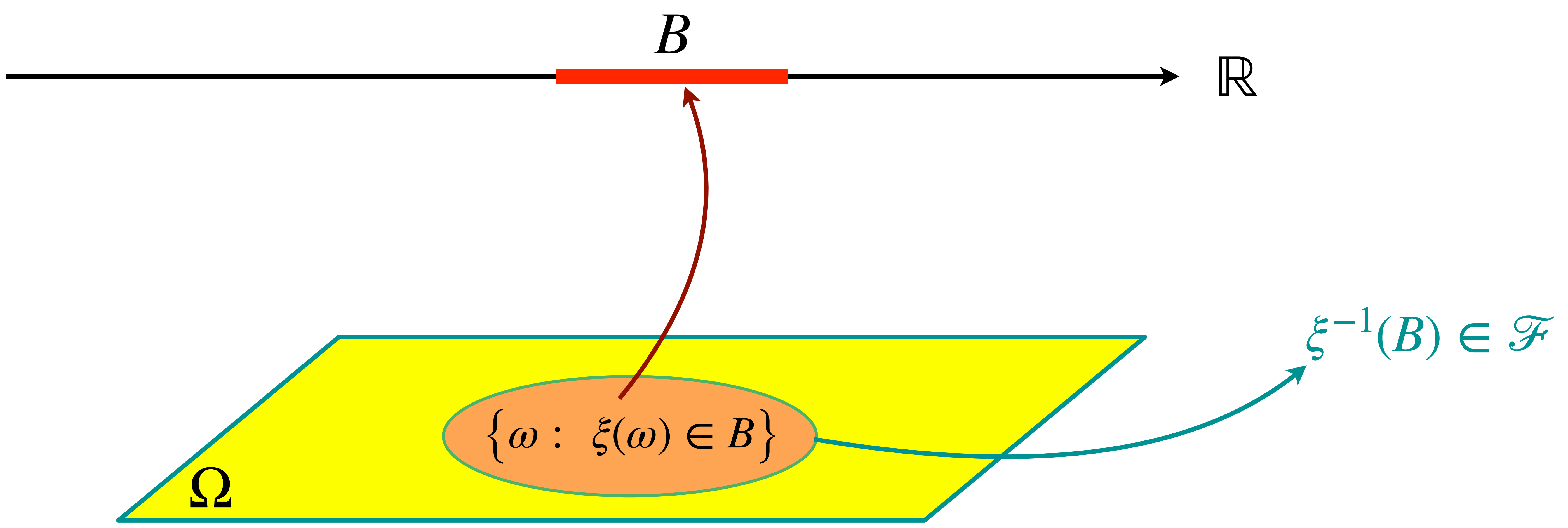
$$\Longrightarrow \quad P(\xi=i)=\frac{1}{6},\quad i=1,2,3,4,5,6$$

一、随机变量的定义

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

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

则称 $\xi(\omega)$ 为随机变量，而 $P\left(\xi(\omega) \in B\right)$ 称为随机变量 $\xi(\omega)$ 的概率分布.



一、随机变量的定义

- 例：对于 $\Omega = \{\omega_1\ ,\ \omega_2\ ,\ \omega_3\}$，构造 σ 代数 $\mathscr{F} = \{\varnothing\ ,\ \{\omega_1\ ,\ \omega_2\}\ ,\ \{\omega_3\}\ ,\ \Omega\}$，

再定义概率 $\begin{cases} P(\varnothing) & = 0 \\ P(\{\omega_1, \omega_2\}) & = \dfrac{2}{3} \\ P(\{\omega_3\}) & = \dfrac{1}{3} \\ P(\Omega) & = 1 \end{cases}$

$(\Omega, \mathscr{F}, P)$
是一个概率空间

▸ 定义实值函数 $\xi(\omega_k) = k, (1 \le k \le 3)$

ξ 不是随机变量

$B = (0\ ,\ 1.5) \implies \{\omega:\ \xi(\omega) \in B\} = \{\omega_1\} \notin \mathscr{F}$

$\mathbb{R}$ 中的 Borel 点集

一、随机变量的定义

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

$$\left\{\omega:\ \xi(\omega) \in B\right\} \in \mathscr{F}$$

则称 $\xi(\omega)$ 为随机变量，而 $P\left(\xi(\omega) \in B\right)$ 称为随机变量 $\xi(\omega)$ 的概率分布.

  - 特别，在上述定义中若取 $B = (-\infty, x)$，则有 $\left\{\omega:\ \xi(\omega) < x\right\} \in \mathscr{F}$.

    从而 $P\left\{\xi(\omega) < x\right\}$ 有定义.

$$\Longrightarrow \quad P\left\{a \le \xi(\omega) < b\right\} = P\left\{\xi(\omega) < b\right\} - P\left\{\xi(\omega) < a\right\}$$

如可列个互不相容左闭右开区间的并

  - $\forall x \in \mathbb{R}$，只要给出 $P\left\{\xi(\omega) < x\right\}$，我们就可以计算出 $\xi(\omega)$ 落入 $\left[a, b\right)$ 内的概率.
  - 再由概率的性质，还可以计算 $\xi(\omega)$ 落入 $\mathbb{R}$ 上更为复杂的点集中的概率.

一、随机变量的定义

- 定义：称

$$F(x) = P\left\{\xi(\omega) \le x\right\}, \quad -\infty < x < \infty$$

为随机变量 $\xi(\omega)$ 的 (累积) 分布函数 (cumulative distribution function, c.d.f.). 记作 $\xi(\omega) \sim F(x)$.

- 根据定义，随机变量是样本点的函数，因此在试验前我们只能知道它可能取哪些值，而不能确知它取何值，这就是随机性. 但试验结束之后，随机变量的取值就确定了，不再有随机性.
- 为了计算概率，必须要求随机变量具有可测性，引进分布函数则将关于随机变量概率的计算化为对分布函数的数值进行运算.
- 由测度论的方法还可以证明：分布函数可以唯一决定概率分布. (测度扩张的唯一性参见《现代概率论基础》，汪嘉冈，P48.)

二、分布函数的性质

① **单调性**：若 $a < b$，则 $F(a) \le F(b)$.

$$F(b) - F(a) = P\{\xi \le b\} - P\{\xi \le a\}$$

$$= P\{a < \xi \le b\}$$

$$\ge 0$$

分布函数单调不降

二、分布函数的性质

① **单调性**：若 $a<b$，则 $F(a)\le F(b)$.

② **有界性**：$0\le F(x)\le 1$，且 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$，$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$.

$$P\{-\infty<\xi<+\infty\}=\sum_{n=-\infty}^{+\infty}P\{n<\xi\le n+1\}=\sum_{n=-\infty}^{+\infty}\left[F(n+1)-F(n)\right]$$

$$=\lim_{n\to+\infty}F(n)-\lim_{m\to-\infty}F(m)=1$$

$F(x)$ 单调有界 $\Longrightarrow$ $\lim\limits_{x\to-\infty}F(x)=\lim\limits_{m\to-\infty}F(m)$ 与 $\lim\limits_{x\to+\infty}F(x)=\lim\limits_{n\to+\infty}F(n)$ 存在

又因为 $0\le F(x)\le 1$ $\Longrightarrow$ $\lim\limits_{x\to-\infty}F(x)=\lim\limits_{m\to-\infty}F(m)=0$

$$\lim_{x\to+\infty}F(x)=\lim_{n\to+\infty}F(n)=1$$

二、分布函数的性质

① 单调性：若 $a<b$，则 $F(a)\le F(b)$.

② 有界性：$0\le F(x)\le 1$，且 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$，$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$.

③ 右连续性：$F(x+0)=\lim\limits_{t\to x+0}F(t)=F(x)$.

由于 $F(x)$ 单调不降、有界，所以只须证明对任一单调下降的序列

$$x_1>x_2>\cdots>x_n>x_0\cdots\ ,\qquad x_n\longrightarrow x_0\quad(n\to\infty)$$

都有 $\lim\limits_{n\to\infty}F(x_n)=F(x_0)$ 成立即可.

$$F(x_1)-F(x_0)=P\{x_0<\xi\le x_1\}=\sum_{n=1}^{\infty}\left[F(x_n)-F(x_{n+1})\right]=F(x_1)-\lim_{n\to\infty}F(x_n)$$

$$\Longrightarrow\quad F(x_0+0)=\lim_{n\to\infty}F(x_n)=F(x_0)$$

- 定义：若 $S_n \in \mathscr{F}$，$n = 1, 2, \ldots$ 且 $S_n \subset S_{n+1}$，则 S_n 是 $\mathscr{F}$ 中的一个单调不减的集序列.

  若 $S_n \in \mathscr{F}$，$n = 1, 2, \ldots$ 且 $S_n \supset S_{n+1}$，则 S_n 是 $\mathscr{F}$ 中的一个单调不增的集序列.

- 定义：对于 $\mathscr{F}$ 上的集合函数 $P(\cdot)$，若它对 $\mathscr{F}$ 中任何一个单调不减的集序列 $\{S_n\}$ 均有

  $\lim\limits_{n\to\infty} P(S_n) = P\left(\lim\limits_{n\to\infty} S_n\right)$ 成立，则称 $P(\cdot)$ 是下连续的.

  对于 $\mathscr{F}$ 上的集合函数 $P(\cdot)$，若它对 $\mathscr{F}$ 中任何一个单调不增的集序列 $\{S_n\}$ 均有

  $\lim\limits_{n\to\infty} P(S_n) = P\left(\lim\limits_{n\to\infty} S_n\right)$ 成立，则称 $P(\cdot)$ 是上连续的.

二、分布函数的性质

① 单调性：若 $a < b$，则 $F(a) \le F(b)$.

② 有界性：$0 \le F(x) \le 1$，且 $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0$，$F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1$.

③ 右连续性：$F(x+0) = \lim\limits_{t\to x+0} F(t) = F(x)$.

由于 $F(x)$ 单调不降、有界，所以只须证明对任一单调下降的序列

$$x_1 > x_2 > \cdots > x_n > \cdots > x_0\,, \qquad x_n \longrightarrow x_0 \quad (n \to \infty)$$

都有 $\lim\limits_{n\to\infty} F(x_n) = F(x_0)$ 成立即可.

亦可利用概率的上连续性来证明

$$1 - F(x_0) = P\{\omega : \ \xi(\omega) > x_0\} = P\left\{\bigcup_{n=1}^{\infty}\{\omega : \ \xi(\omega) > x_n\}\right\}$$

$$= 1 - P\left\{\bigcap_{n=1}^{\infty}\{\omega : \xi(\omega) \le x_n\}\right\} = 1 - \lim_{n\to\infty} P\{\omega : \ \xi(\omega) \le x_n\} = 1 - \lim_{n\to\infty} F(x_n)$$

二、分布函数的性质

- 有了分布函数，关于随机变量的许多概率都能方便算出：

① $$P\{\xi(\omega)\le a\} = P\left\{\bigcap_{n=1}^{\infty}\left\{\xi(\omega)\le a+\frac{1}{n}\right\}\right\} = \lim_{n\to+\infty} P\left\{\xi(\omega)\le a+\frac{1}{n}\right\} = \lim_{n\to+\infty} F\left(a+\frac{1}{n}\right) = F(a+0)$$

② $$P\{\xi(\omega)=a\} = P\{\xi(\omega)\le a\} - P\{\xi(\omega)<a\} = F(a) - F(a-0)$$

③ $$P\{\xi(\omega)\ge a\} = 1 - P\{\xi(\omega)<a\} = 1 - F(a-0)$$

④ $$P\{\xi(\omega)>a\} = P\{\xi(\omega)\ge a\} - P\{\xi(\omega)=a\} = 1 - F(a-0) - F(a) + F(a-0) = 1 - F(a)$$

⑤ $$P\{a<\xi(\omega)\le b\} = P\{\xi(\omega)\le b\} - P\{\xi(\omega)\le a\} = F(b) - F(a)$$

三、离散型随机变量

(discrete random variable)

- 定义：若随机变量 ξ 的全部可能取值是有限个或可列无限多个，则称 ξ 是离散型随机变量.
- 定义：设离散型随机变量 ξ 的所有可能取值为 x_k , $k = 1, 2, \ldots$, 则称事件 $\{\xi = x_k\}$ 的概率

  $P\{\xi = x_k\} \triangleq p_k$, $k = 1, 2, \ldots$ 为随机变量 ξ 的概率分布 (probability distribution) 或分布律.

  - 分布列：概率分布的表格形式

ξ	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

  - 概率分布的性质：

    ① $p_k \geq 0$, $k = 1, 2, \ldots$

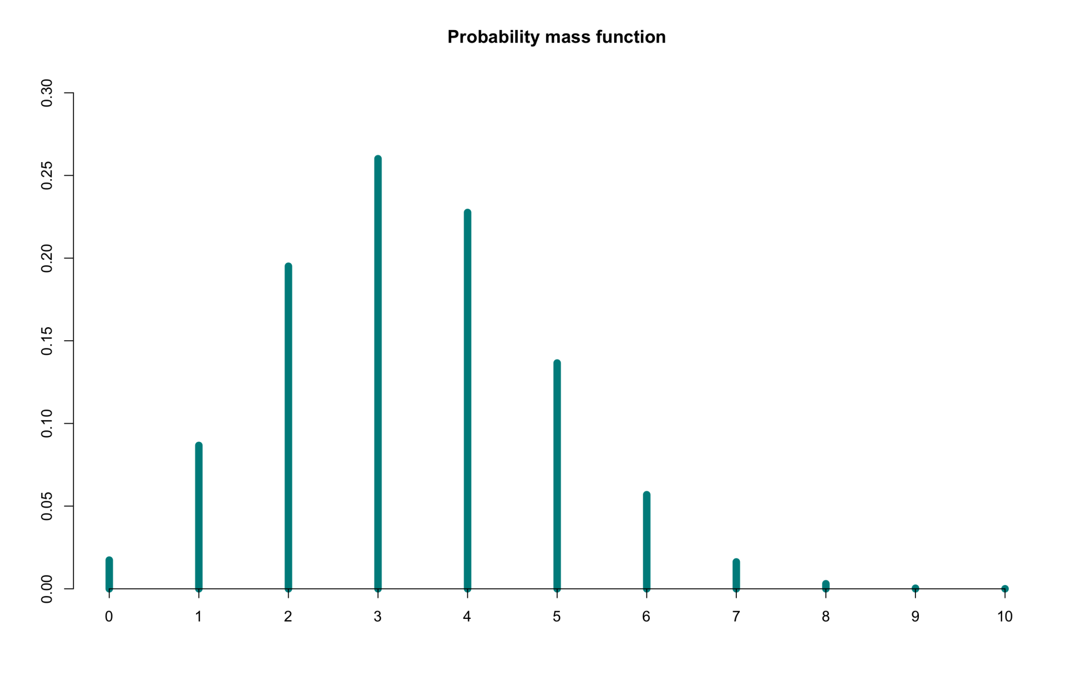
    ② $\sum_{k=1}^{\infty} p_k = 1$

三、离散型随机变量

- 已知分布列求分布函数：$F(x) = P\left\{\xi \le x\right\} = \sum\limits_{x_k \le x} p_k\ ,\qquad \forall x \in \mathbb{R}$
- 已知分布函数求分布列：$P\left\{\xi = x\right\} = F(x) - F(x-0)\ ,\qquad \forall x \in \mathbb{R}$

```
## 二项分布 b(10, 1/3) 的分布列
n = 10
p = dbinom(0:n, n, p = 1/3)
probs = data.frame(n = 0:n, p = p)
round(probs, digits = 4)
plot(probs, type = 'h', axes = FALSE, xlab = '', ylab = '', ylim = c(0, 0.3), lwd = 8,
   col = 'cyan4', main = 'Probability mass function')
axis(1, at = 0:n, pos = 0)
axis(2)
```

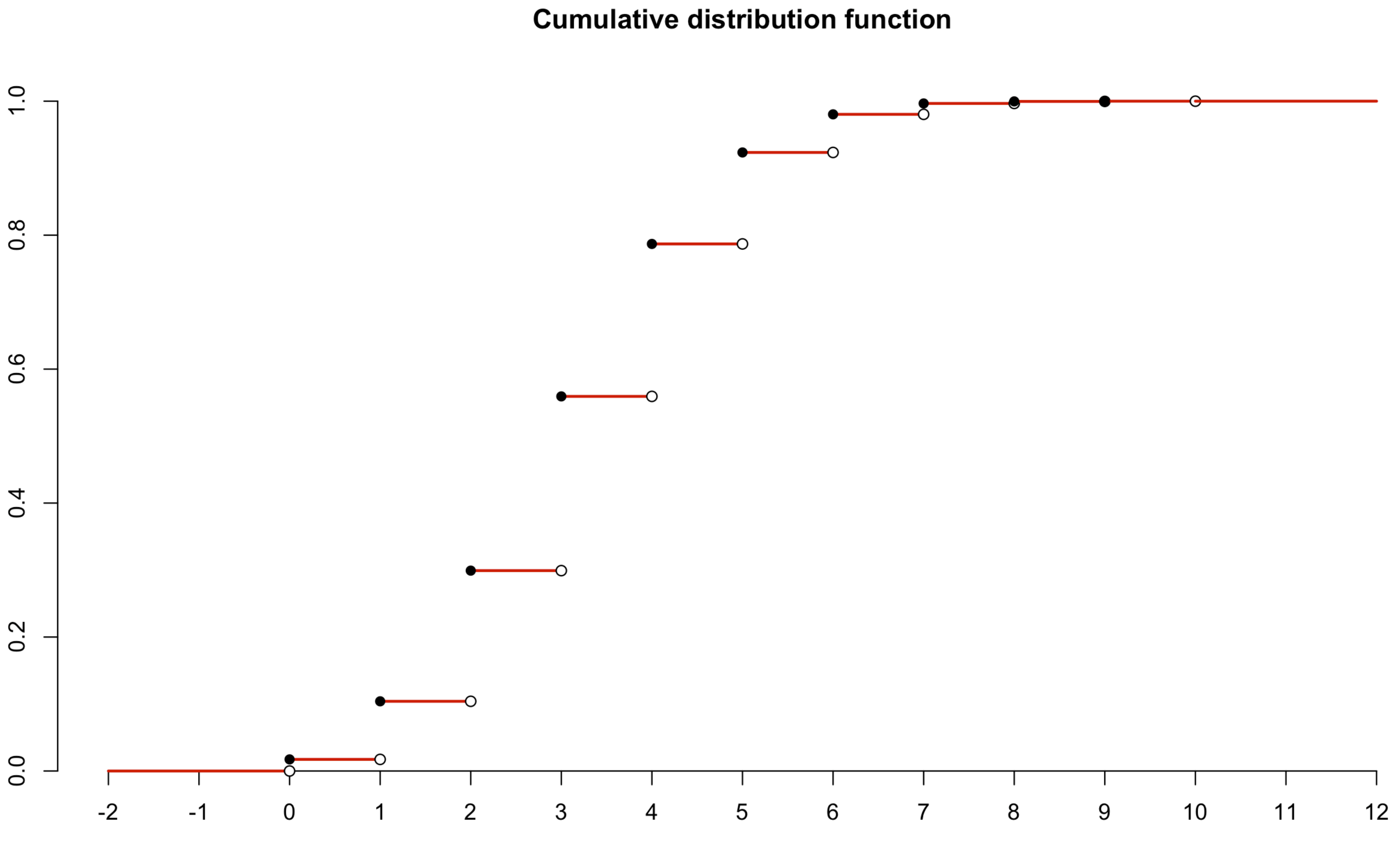
```
> round(probs, digits = 4)
    n      p
1   0 0.0173
2   1 0.0867
3   2 0.1951
4   3 0.2601
5   4 0.2276
6   5 0.1366
7   6 0.0569
8   7 0.0163
9   8 0.0030
10  9 0.0003
11 10 0.0000
```



三、离散型随机变量

- 已知分布列求分布函数：$F(x)=P\{\xi\le x\}=\sum\limits_{x_k\le x}p_k$ ，　$\forall x\in\mathbb{R}$
- 已知分布函数求分布列：$P\{\xi=x\}=F(x)-F(x-0)$ ，　$\forall x\in\mathbb{R}$

```
## 二项分布
n = 10
P = pbinom(0
cdf = data.fra
round(cdf, dig
plot(cdf, type
    main = 'Cu
axis(1, at = -2
axis(2)
lines(c(-2, 0),
for (i in 0:(n-1
  lines(c(i, i+1
  points(i, P[i-
  points(i+1, P
}
points(0, 0, p
lines(c(10, 12
```




```
> round(cdf, digits = 4)
    n      p
1   0 0.0173
2   1 0.1040
3   2 0.2991
4   3 0.5593
5   4 0.7869
6   5 0.9234
7   6 0.9803
8   7 0.9966
9   8 0.9996
10  9 1.0000
11 10 1.0000
```

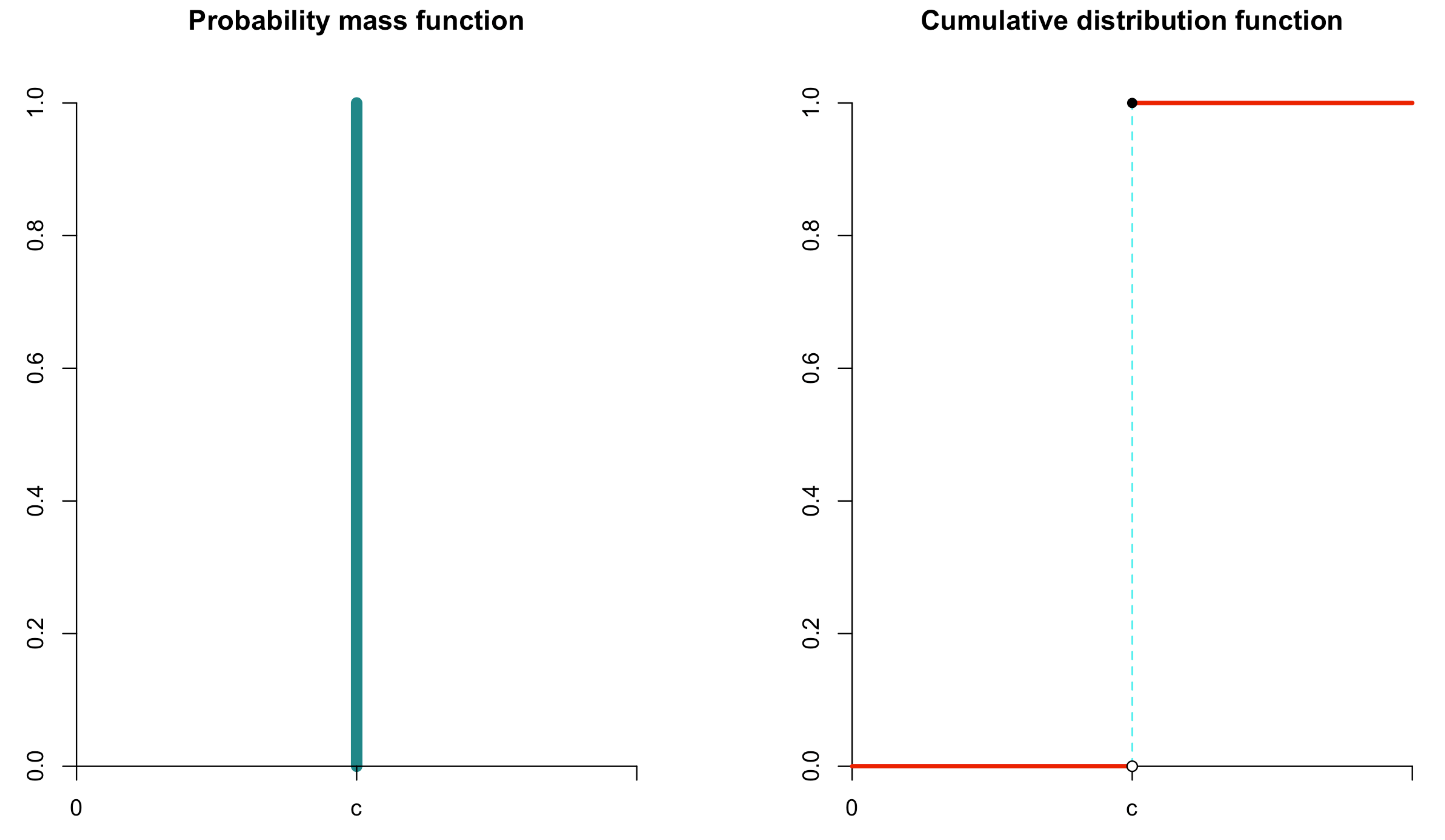
三、离散型随机变量

- 常见的离散型随机变量及其分布

① 退化 (degene

- 分布律：
- 分布函

```
## 单点分布
par(mfrow = c(1, 2))
x = c(0, 2)
y = c(0, 1)
plot(x, y, type = 'n', xlab = '', ylab = '', axes = FA
lines(c(1, 1), c(0, 1), lwd = 8, col = 'cyan4')
axis(1, pos = 0, at = 0:2, labels = c(0, 'c', ''))
axis(2, pos = 0)
plot(x, y, type = 'n', xlab = '', ylab = '', axes = FA
axis(1, pos = 0, at = 0:2, labels = c(0, 'c', ''))
axis(2, pos = 0)
lines(c(0, 1), c(0, 0), lwd = 3, col = 'red2')
lines(c(1, 2), c(1, 1), lwd = 3, col = 'red2')
lines(c(1, 1), c(0, 1), lwd = 1, col = 'cyan2', lty = 2)
points(1, 1, pch = 21, bg = 'white')
points(1, 0, pch = 16)
```

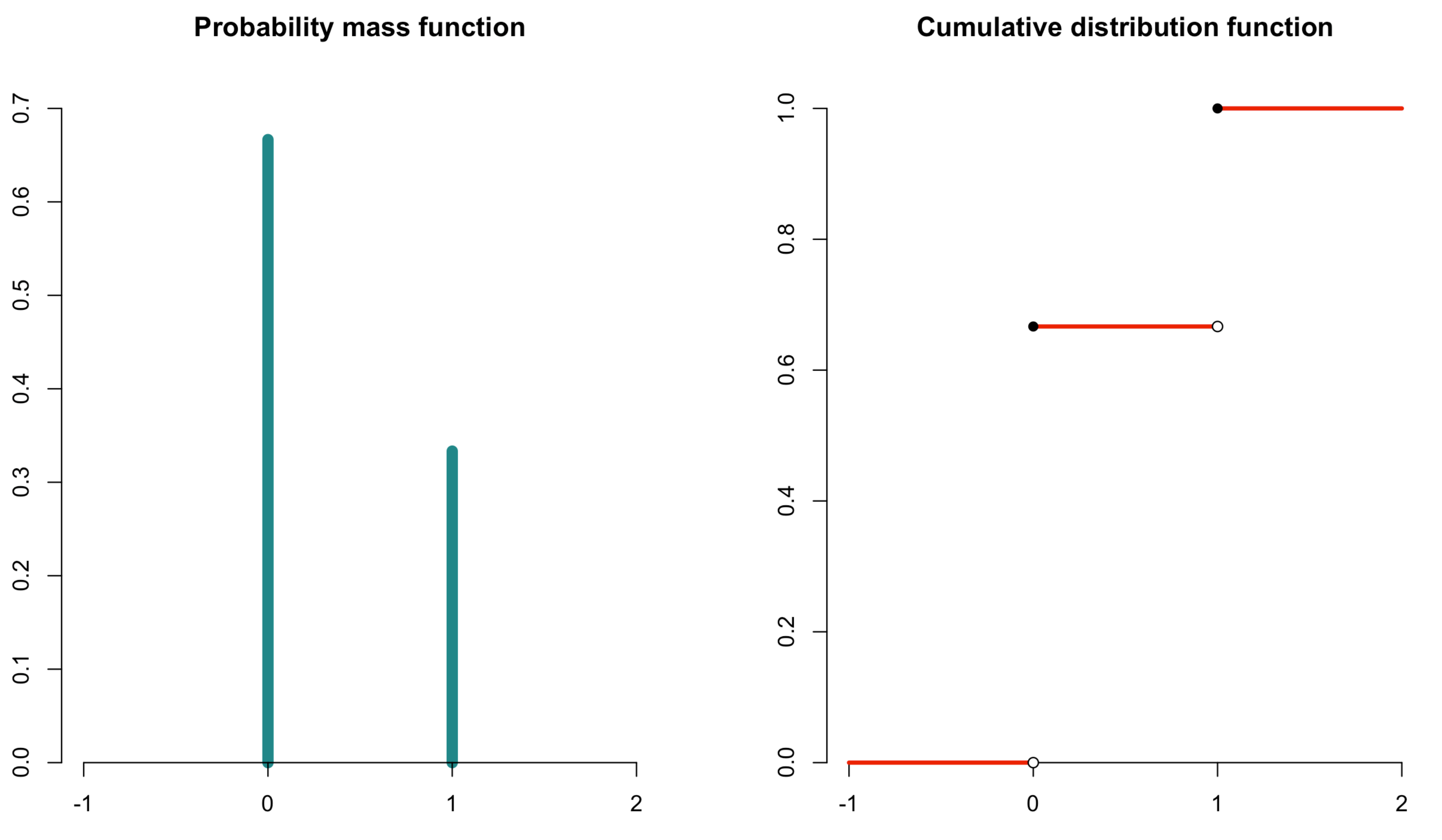


三、离散型随机变量

- 常见的离

② Ber

▶



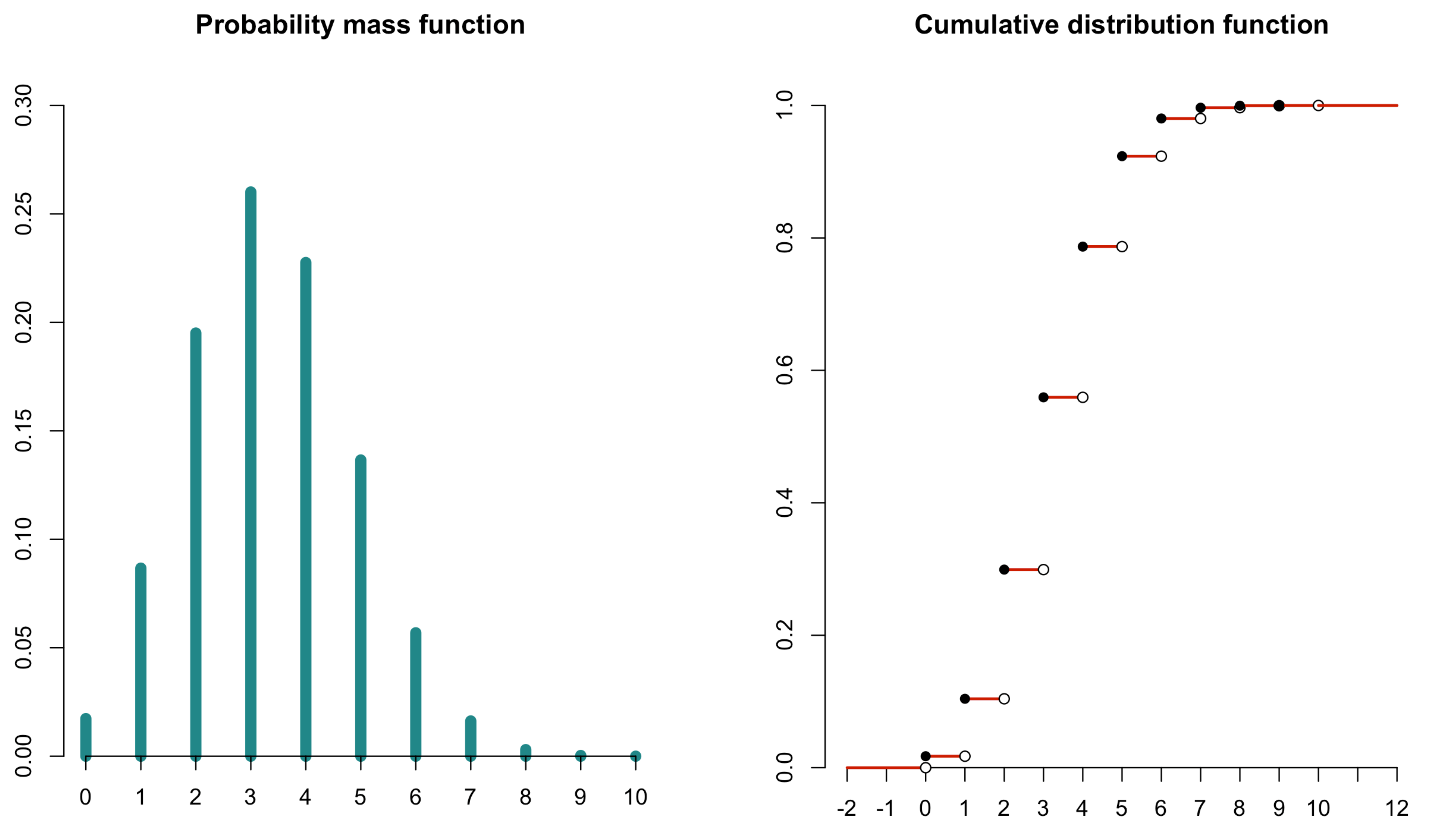
0	1
$1-p$	p

```
## Bernoulli 分布
par(mfrow = c(1, 2))
n = 1
plot(0:n, dbinom(0:n, 1, 1/3), type
axis(1, pos = 0, at = -1:2)
axis(2)
plot(-1:2, rep(1, 4), type = 'n', xlab
axis(1, pos = 0, at = -1:2)
axis(2)
lines(c(-1, 0), c(0, 0), lwd = 3, col
lines(c(0, 1), c(pbinom(0, 1, 1/3),
lines(c(1, 2), c(1, 1), lwd = 3, col =
points(0, pbinom(0, 1, 1/3), pch =
points(1, 1, pch = 21, bg = 'white')
points(1, pbinom(0, 1, 1/3), pch = 16)
points(0, 0, pch = 16)
```

三、离散型随机变量

- 常见的离

③ 二项



$B(n, p)$

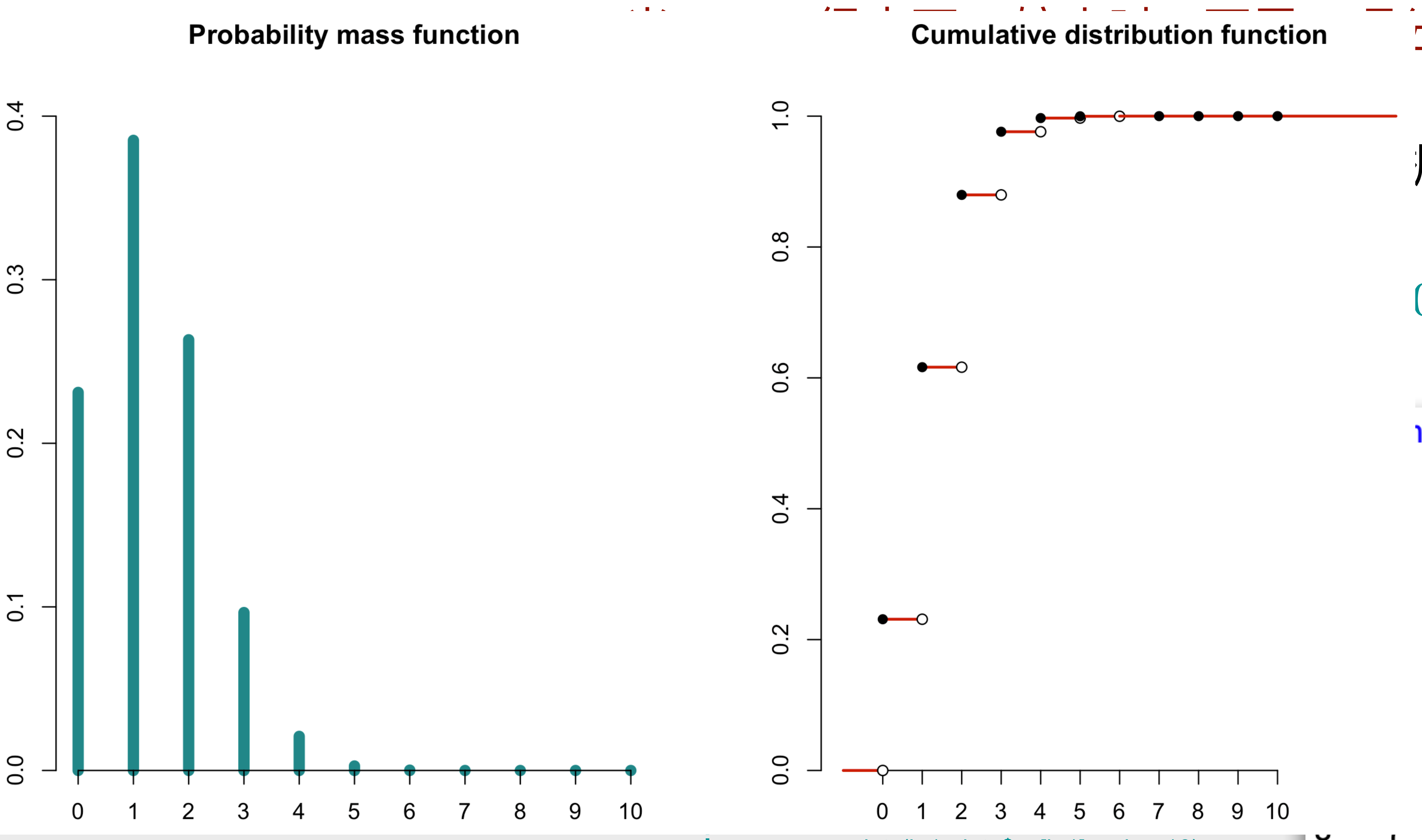
```
## 二项分布
par(mfrow = c(1, 2))
n = 10
p = dbinom(0:n, n, p = 1/3)
probs = data.frame(n = 0:n, p = p
round(probs, digits = 4)
plot(probs, type = 'h', axes = FAL
axis(1, at = 0:n, pos = 0)
axis(2)
P = pbinom(0:n, n, prob = 1/3)
cdf = data.frame(n = 0:n, p = P)
round(cdf, digits = 4)
plot(cdf, type = 'n', axes = FALSE
axis(1, at = -2:12, pos = 0)
axis(2)
lines(c(-2, 0), c(0, 0), col = 'red3',
for (i in 0:(n-1)) {
  lines(c(i, i+1), c(P[i+1], P[i+1]), col = 'red3', lwd = 2)
  points(i, P[i+1], pch = 21, bg = 'white')
  points(i+1, P[i+1], pch = 16)
}
points(0, 0, pch = 16)
lines(c(10, 12), c(1, 1), col = 'red3', lwd = 2)
```

三、离散型随机变量

- 常见的离散 … 分布近似

④ 超几 … 几取 k 个球出

仅当 $\max\{0, k-n\}$ … $0, 1, 2, \ldots, m+n\}$




```
## 超几何分布
par(mfrow = c(1, 2))
m = 10
n = 90
k = 13
x = 0:min(k, m)
probs = dhyper(x, m, n, k)
cumprobs = phyper(x, m, n, k)
data = data.frame(x = x, p = probs,
round(data, digits = 4)

plot(data$x, data$p, type = 'h', axes = FALSE, xlab = '', ylab = '', ylim = c(0, 0.4), lwd = 8,
        col = 'cyan4', main = 'Probability mass function')
axis(1, at = 0:min(k, m), pos = 0)
axis(2)
```

```
  points(i+1, data$cp[i+1], pch = 16)
}
points(0, 0, pch = 16)
lines(c(6, 13), c(1, 1), col = 'red3', lwd = 2)
points(7:10, c(1, 1, 1, 1), pch = 21, bg = 'white
```

```
nd(data, digits = 4)
         p      cp
   0.2311 0.2311
   0.3852 0.6163
   0.2633 0.8796
   0.0965 0.9762
   0.0209 0.9970
   0.0027 0.9998
   0.0002 1.0000
   0.0000 1.0000
9   8 0.0000 1.0000
10  9 0.0000 1.0000
11 10 0.0000 1.0000
```

三、离散型随机变量

- 常见的离

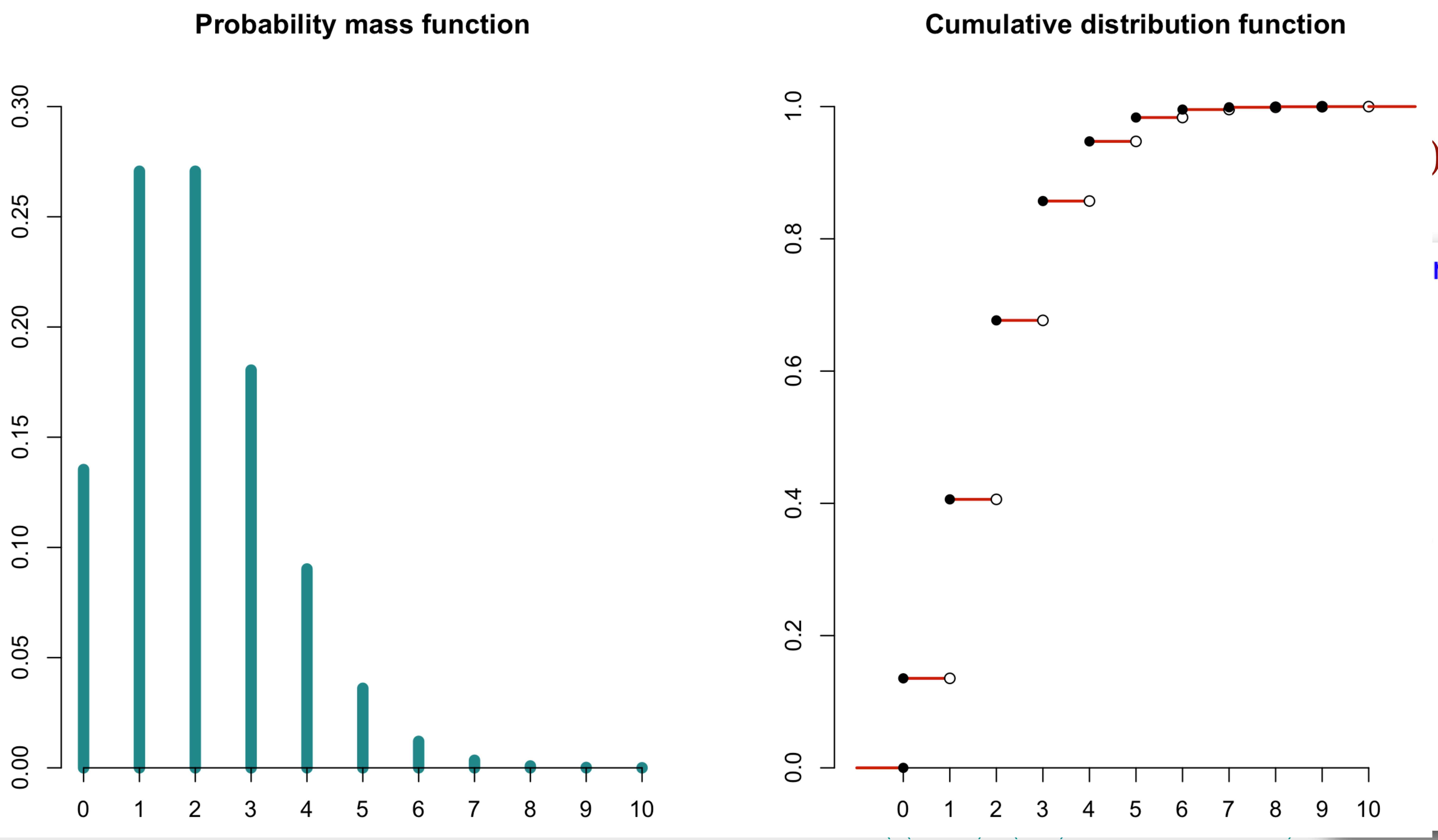
⑤ Pois

)

```
## Poisson 分布
par(mfrow = c(1, 2))
lambda = 2
n = 10
x = 0 : n

probs = dpois(x, lambda)
cumprobs = ppois(x, lambda)
data = data.frame(x = x, p = probs
round(data, digits = 4)

plot(data$x, data$p, type = 'h', axe
          col = 'cyan4', main = 'P
axis(1, at = 0:n, pos = 0)
axis(2)
```




```
nd(data, digits = 4)
     p     cp
0.1353 0.1353
0.2707 0.4060
0.2707 0.6767
0.1804 0.8571
0.0902 0.9473
0.0361 0.9834
0.0120 0.9955
0.0034 0.9989
0.0009 0.9998
0.0002 1.0000
0.0000 1.0000
```

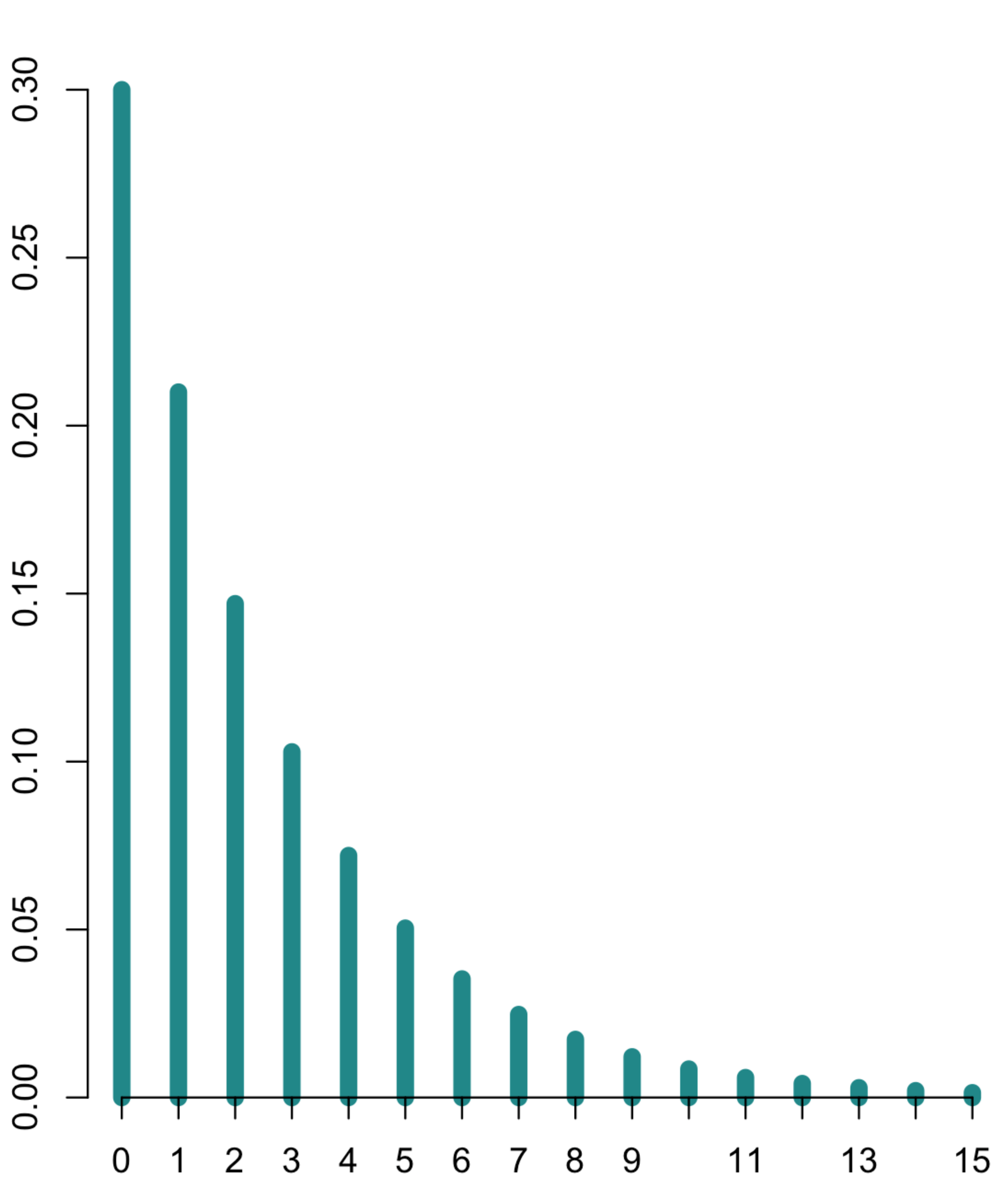
三、离散型随机变量

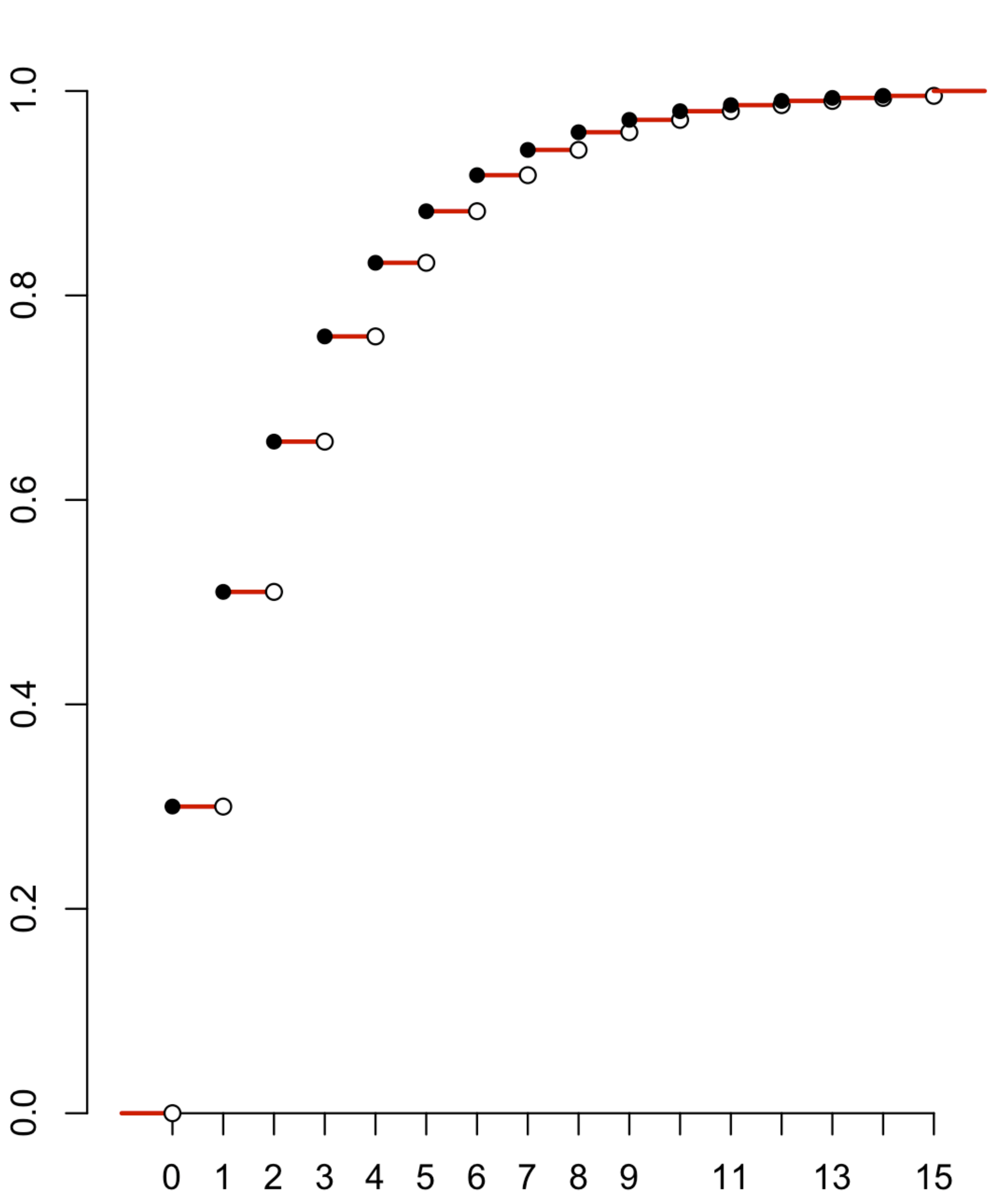
- 常见的离散型随机变量及其分布

⑥ 几何

```
## 几何分布
par(mfrow = c(1, 2))
x = 0 : 15
probs = dgeom(x, prob = 0.3)
cumprobs = pgeom(x, prob = 0.3)
data = data.frame(x = x, p = probs
round(data, digits = 3)

plot(data$x, data$p, type = 'h', axe
            col = 'cyan4', main = 'P
axis(1, at = 0:15, pos = 0)
axis(2)
```






```
nd(data, digits = 3)
          p    cp
      0.300 0.300
      0.210 0.510
      0.147 0.657
      0.103 0.760
      0.072 0.832
      0.050 0.882
      0.035 0.918
      0.025 0.942
      0.017 0.960
      0.012 0.972
      0.008 0.980
      0.006 0.986
      0.004 0.990
14 13 0.003 0.993
15 14 0.002 0.995
16 15 0.001 0.997
```

三、离散型随机变量

- 常见的离散型随机变量及其分布

⑥ 几何 (geometric) 分布：Bernoulli 试验中，事件 A 首次出现时的试验次数.

$$g(x;\ p) = P\left\{\eta = x\right\} = (1-p)^{x-1}p\ ,\quad x = 1,\ 2,\ \ldots$$

▸ 几何分布的无记忆性

定理：取值于自然数的随机变量 η 服从几何分布，当且仅当 η 具有无记忆性

$$P\left\{\eta > m+n \,\middle|\, \eta > m\right\} = P\left\{\eta > n\right\}\ ,\quad \forall\ m,\ n \geq 1$$

或者

$$P\left\{\eta = m+n \,\middle|\, \eta > m\right\} = P\left\{\eta = n\right\}\ ,\quad \forall\ m,\ n \geq 1$$

Remark：几何分布是唯一无记忆的离散型分布.

三、离散型随机变量

- 证明： $\Longrightarrow$ 必要性. 因为 η 服从几何分布，所以

$$P\{\eta > n\} = \sum_{x=n+1}^{\infty} q^{x-1}p = \sum_{x-n-1=0}^{\infty} q^{x-n-1}q^n p$$

$q = 1 - p$

$$= \left(\sum_{k=0}^{\infty} q^k\right) q^n p = \frac{pq^n}{1-q} = q^n, \quad n = 1, 2, \ldots$$

$$\xrightarrow{\forall\, m, n \geq 1} \quad P\left\{\eta > m+n \,\middle|\, \eta > m\right\} = \frac{P\{\eta > m+n\}}{P\{\eta > m\}} = \frac{q^{m+n}}{q^m} = q^n = P\{\eta > n\}$$

因而 η 具体无记忆性.

▸ 几何分布的无记忆性

定理：取值于自然数的随机变量 η 服从几何分布，当且仅当 η 具有无记忆性

$$P\left\{\eta > m+n \,\middle|\, \eta > m\right\} = P\{\eta > n\}, \quad \forall\, m, n \geq 1$$

或者

$$P\left\{\eta = m+n \,\middle|\, \eta > m\right\} = P\{\eta = n\}, \quad \forall\, m, n \geq 1$$

三、离散型随机变量

▸ 几何分布的无记忆性

定理：取值于自然数的随机变量 η 服从几何分布，当且仅当 η 具有无记忆性

$$P\left\{\eta > m+n \middle| \eta > m\right\} = P\{\eta > n\}\ ,\quad \forall\, m, n \geq 1$$

或者

$$P\left\{\eta = m+n \middle| \eta > m\right\} = P\{\eta = n\}\ ,\quad \forall\, m, n \geq 1$$

- 证明：$\Longleftarrow$ 充分性．假设 $Q_n = P\{\eta > n\} > 0\ ,\quad \forall\, n \geq 1$

$$\xrightarrow{\text{乘法公式}} P\{\eta > m+n\} = P\{\eta > m\} \cdot P\left\{\eta > m+n \middle| \eta > m\right\}$$

$$\xrightarrow{\eta\text{ 无记忆}} P\left\{\eta > m+n \middle| \eta > m\right\} = P\{\eta > n\}$$

$$P\{\eta > m+n\} = P\{\eta > m\} \cdot P\{\eta > n\}$$

$$Q_{m+n} = Q_m \cdot Q_n\ ,\quad \forall\, m, n \geq 1$$

$$q \triangleq Q_1$$

$$Q_m = Q_{m-1+1} = Q_{m-1} \cdot Q_1 = Q_{m-2+1} \cdot Q_1 = Q_{m-2} \cdot Q_1^2 = \cdots = Q_1^m = q^m$$

$$P\{\eta = k\} = P\{\eta > k-1\} - P\{\eta > k\} = Q_{k-1} - Q_k = q^{k-1} - q^k = q^{k-1}(1-q) = q^{k-1}p$$

因而 η 服从几何分布．