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
  - 事件独立性：两个事件 $P(AB) = P(A) \cdot P(B)$. 三个事件 $\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$
  - 试验独立性：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.
  - Bernoulli 试验 E：概率空间 $\left(\Omega, \mathscr{F}, P\right)$，其中

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

- Bernoulli 分布
- 二项分布
- 几何 分布
- Pascal 分布
- 多项分布
- Poisson 分布

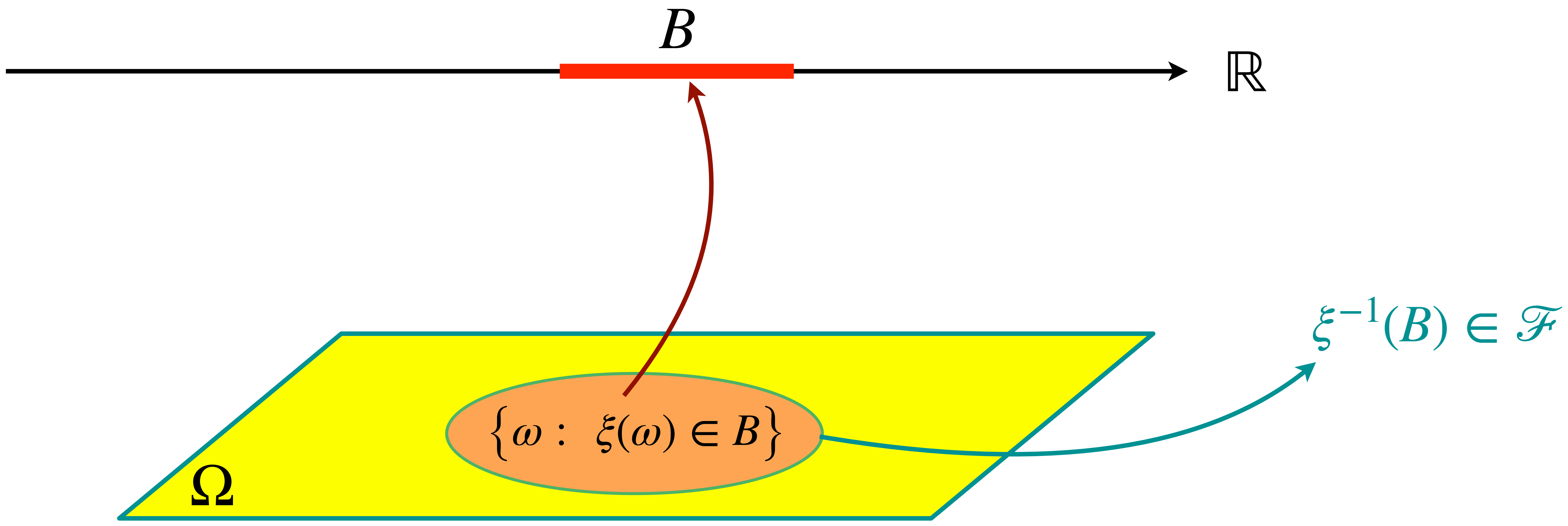
已学知识点

- 第三章 随机变量与分布函数
  - 随机变量：设 $\xi(\omega)$ 是定义于概率空间 $\big(\Omega, \mathscr{F}, P\big)$ 上的单值实函数，如果对于直线上任一

    Borel 点集 B，有

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

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



已学知识点

- 第三章 随机变量与分布函数
  - 分布函数：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 $\xi(\omega) \sim F(x)$.

  - 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：$F(x+0) = \lim\limits_{t\to x+0} F(t) = F(x)$.

④ 分布函数可以唯一确定概率分布.

已学知识点

- 第三章 随机变量与分布函数

  - 分布函数：称

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

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

  - 利用分布函数计算概率

① $P\left\{\xi(\omega) \leq a\right\} = F(a)$

② $P\left\{\xi(\omega) = a\right\} = F(a) - F(a-0)$

③ $P\left\{\xi(\omega) \geq a\right\} = 1 - F(a-0)$

④ $P\left\{\xi(\omega) > a\right\} = 1 - F(a)$

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

已学知识点

- 第三章 随机变量与分布函数
  - 离散型随机变量：随机变量 ξ 的全部可能取值是有限个或可列无限多个
    - 概率分布 (分布律)：ξ 的所有可能值 x_k，$k = 1, 2, \ldots$，则其概率分布 $P\{\xi = x_k\} \triangleq p_k$，$k = 1, 2, \ldots$
    - 分布列：

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 $\sum_{k=1}^{\infty} p_k = 1$

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

已学知识点

- 第三章 随机变量与分布函数
  - 常见的离散型随机变量及其分布

① 退化 (degenerate) 分布：$P\left\{\xi = c\right\} = 1$

② Bernoulli 分布：$P\left\{\xi = k\right\} = b(k;\ 1, p) = p^k(1-p)^{1-k}\,, \quad k = 0,\ 1$

③ 二项 (binomial) 分布：$P\left\{\xi = k\right\} = b(k;\ n, p) = \binom{n}{k} p^k(1-p)^{n-k}\,, \quad k = 0,\ 1,\ 2,\ \dots,\ n$

- 二项分布可以分解为 n 个独立同分布的 Bernoulli 分布的随机变量之和.

④ 超几何 (hyper-geometric) 分布：$h_x = P\left\{\xi = x\right\} = \dfrac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}\,, \quad x = 0,\ 1,\ 2,\ \dots,\ k$

⑤ Poisson 分布：$P\left\{\xi = x\right\} = \dfrac{\lambda^x}{x!}\mathrm{e}^{-\lambda}\,, \quad x = 0,\ 1,\ 2,\ \dots$

已学知识点

- 第三章 随机变量与分布函数
  - 常见的离散型随机变量及其分布

⑥ 几何 (geometric) 分布：$g(x,\ p) = P\left\{\eta = x\right\} = (1-p)^{x-1}p\ ,\quad x = 1,\ 2,\ \ldots$

- 几何分布的无记忆性：$P\left\{\eta > m+n \,\middle|\, \eta > m\right\} = P\left\{\eta > n\right\}\ ,\quad \forall\ m,\ n \geq 1$

⑦ Pascal 分布：$P\left\{\zeta = x\right\} = \binom{x-1}{r-1} p^r q^{x-r}\ ,\quad x = r,\ r+1,\ \ldots$

- Pascal 分布可以分解为 r 个独立同分布的几何分布的随机变量之和.

⑧ 负二项分布：$Nb(x;\ r,\ p) = \binom{-r}{x} p^r\,(-q)^x\ ,\quad x = 0,\ 1,\ 2,\ \ldots$

已学知识点

- 第三章 随机变量与分布函数
  - 连续型随机变量：存在非负可积函数 $p(x)$，使得 $F(x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t$
    - 其分布函数 $F(x)$ 绝对连续、几乎处处可导：$p(x) = F'(x) = \dfrac{\mathrm{d}}{\mathrm{d}x}F(x)$, for a.e. $x \in \mathbb{R}$
    - 概率密度的性质：

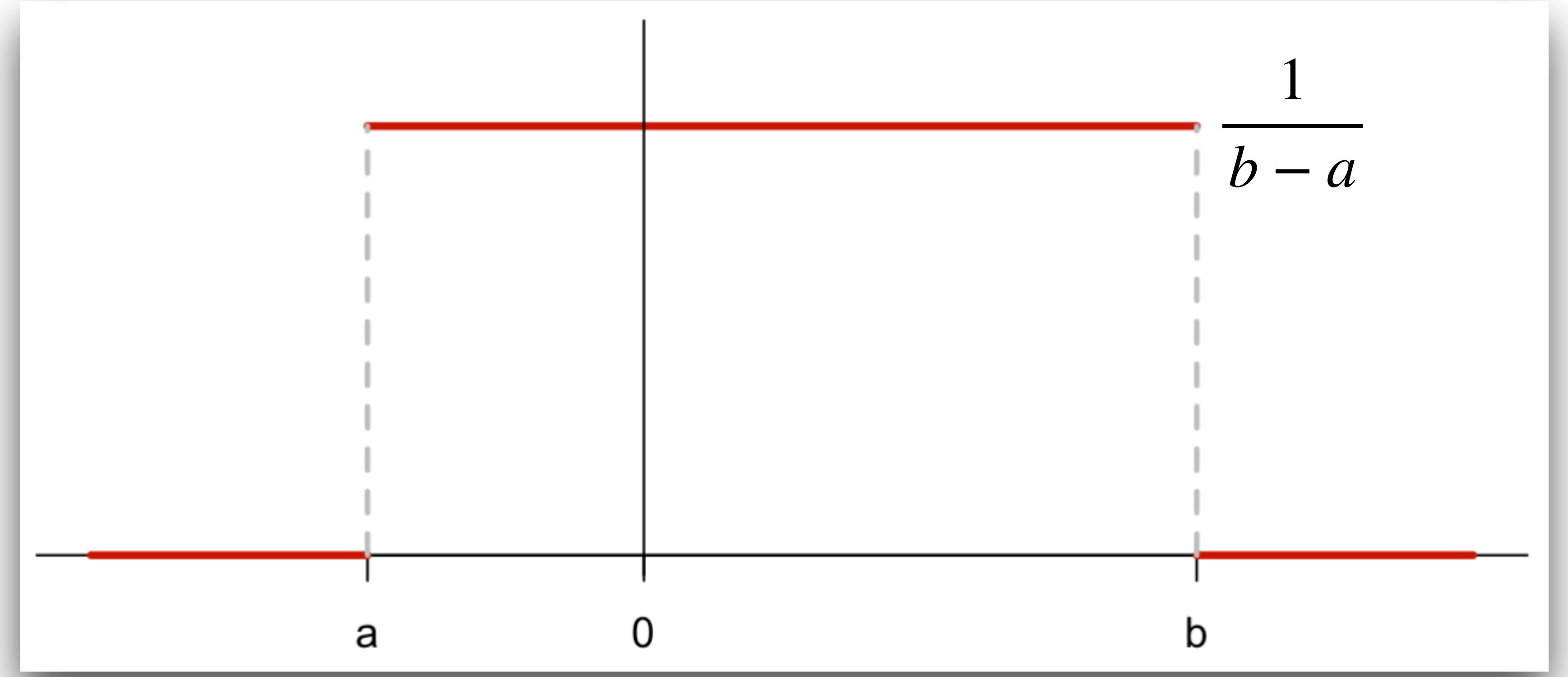
      ① $p(x) \geq 0$　　② $\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1$
    - 概率计算：对 $\mathbb{R}$ 中任一 Borel 点集 B 有 $P\{\xi \in B\} = \int_B p(x)\,\mathrm{d}x$.

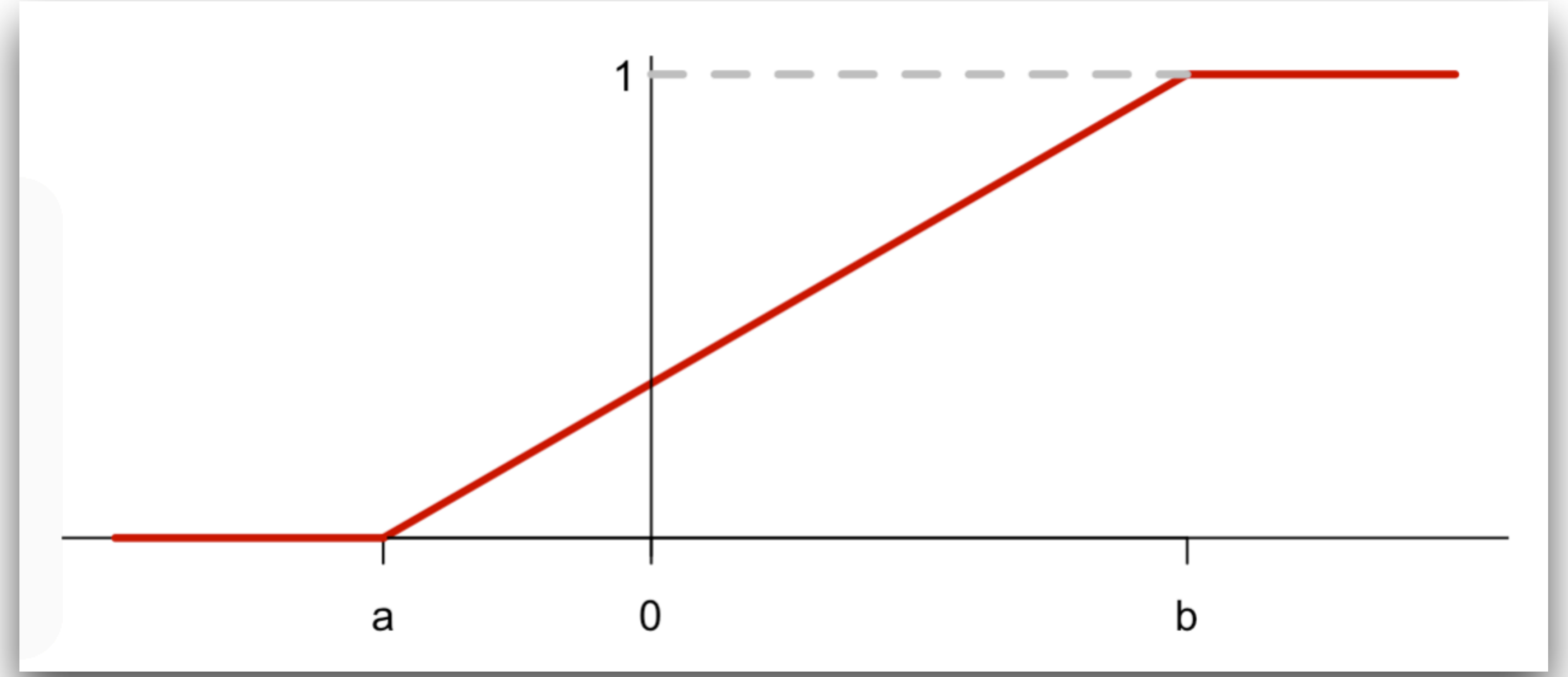
已学知识点

- 第三章 随机变量与分布函数
  - 常见的连续型随机变量及其分布

① 均匀 (uniform) 分布：$\xi \sim U[a, b]$. p.d.f. $p(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{其它} \end{cases}$

c.d.f. $F(x) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ 1, & x > b \end{cases}$





已学知识点

- 第三章 随机变量与分布函数
  - 常见的连续型随机变量及其分布

② 正态 (normal) 分布：$\xi \sim N\left(\mu,\ \sigma^2\right)$.

p.d.f. $p(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$

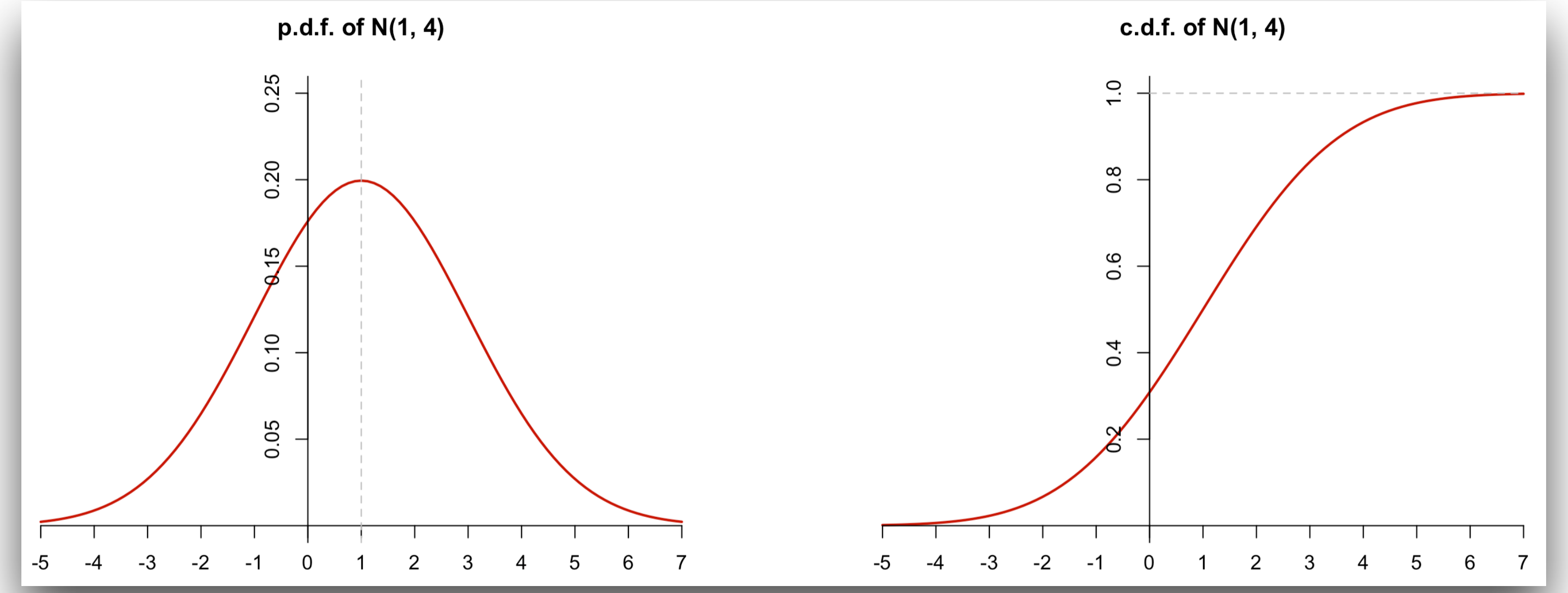
c.d.f. $F(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\ \sigma}\, \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\, \mathrm{d}t, \quad -\infty < x < +\infty$

标准正态分布：$N(0,\ 1)$

p.d.f. $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{x^2}{2}}, \quad -\infty < x < +\infty$

c.d.f. $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\, \mathrm{d}t, \quad -\infty < x < +\infty$

▸ 性质：$\xi \sim N\left(\mu,\ \sigma^2\right) \implies \dfrac{\xi - \mu}{\sigma} \sim N(0,\ 1)$.



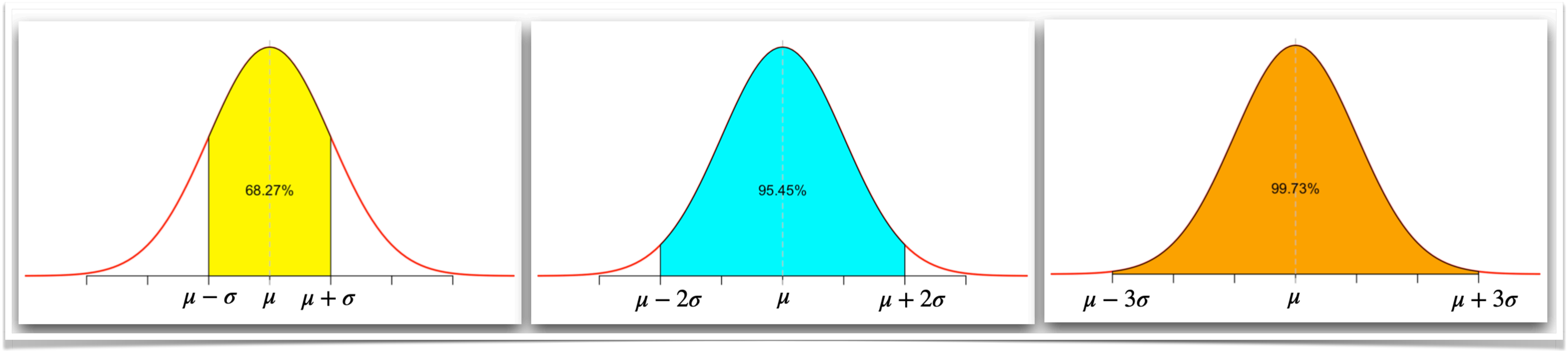
已学知识点

- 第三章 随机变量与分布函数
  - 常见的连续型随机变量及其分布

② 正态 (normal) 分布：$\xi \sim N\left(\mu,\, \sigma^2\right)$.

p.d.f. $p(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$

c.d.f. $F(x) = \displaystyle\int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt, \quad -\infty < x < +\infty$



四、连续型随机变量

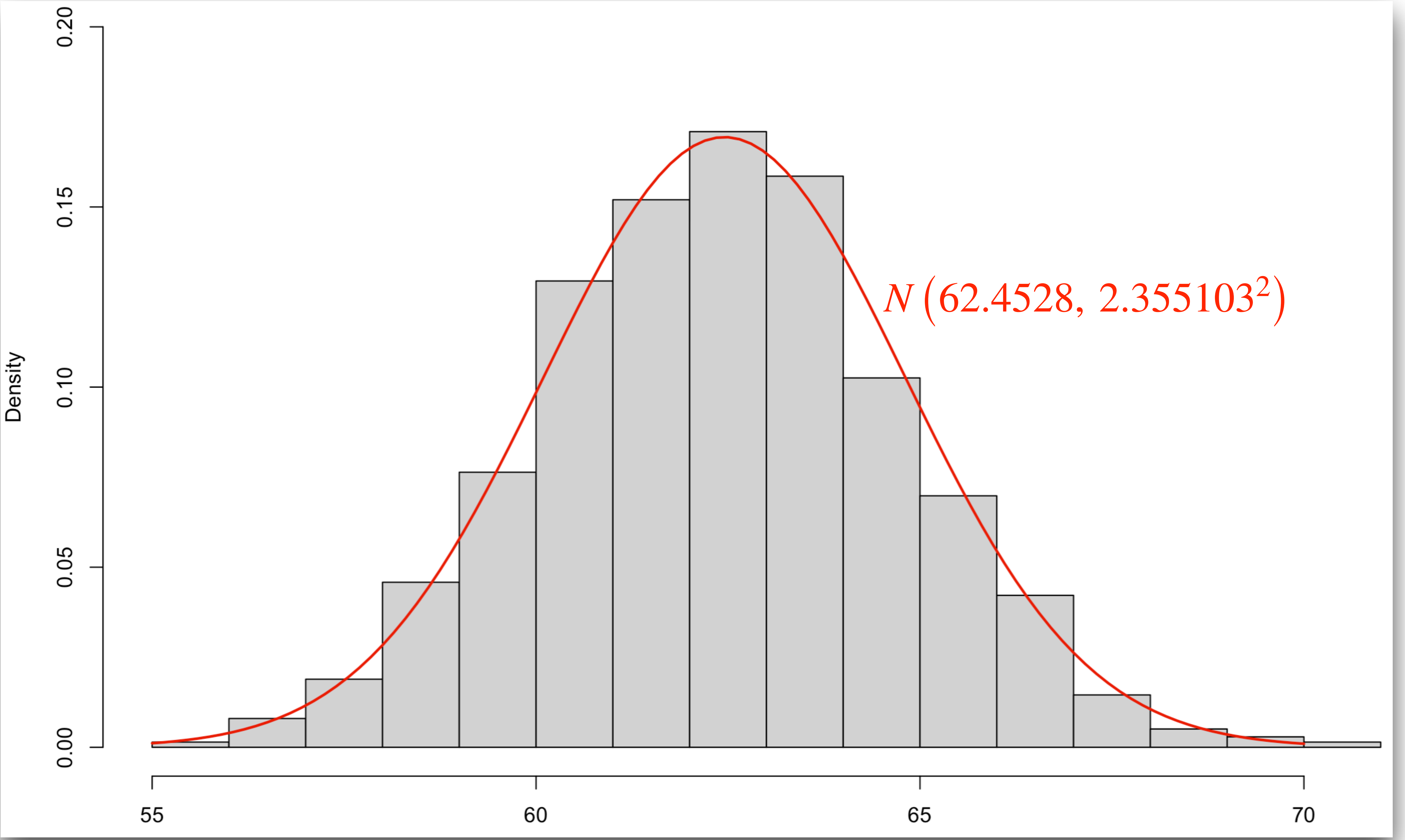
- 常见的连续型随机变量及其分布

② 正态 (normal) 分布：$\xi \sim N\left(\mu, \sigma^2\right)$. 是概率论中最重要的分布

▶ 例：$\mathfrak{R}$ 包 alr4 中的数据集 Heights 的第一个变量 mheight 给出了 1375 位母亲的身高数据 (单位：英尺).

```
library(alr4)
x = Heights$mheight
hist(x, freq = FALSE, ylim = c(0, 0.20), xlab = '', main = '')
(mu = mean(x))
(sigma = sd(x))
curve(dnorm(x, mu, sigma), 55, 70, col = 'red2', add = TRUE, lwd = 2)
```

```
> (mu = mean(x))
[1] 62.4528
> (sigma = sd(x))
[1] 2.355103
```



四、连续型随机变量

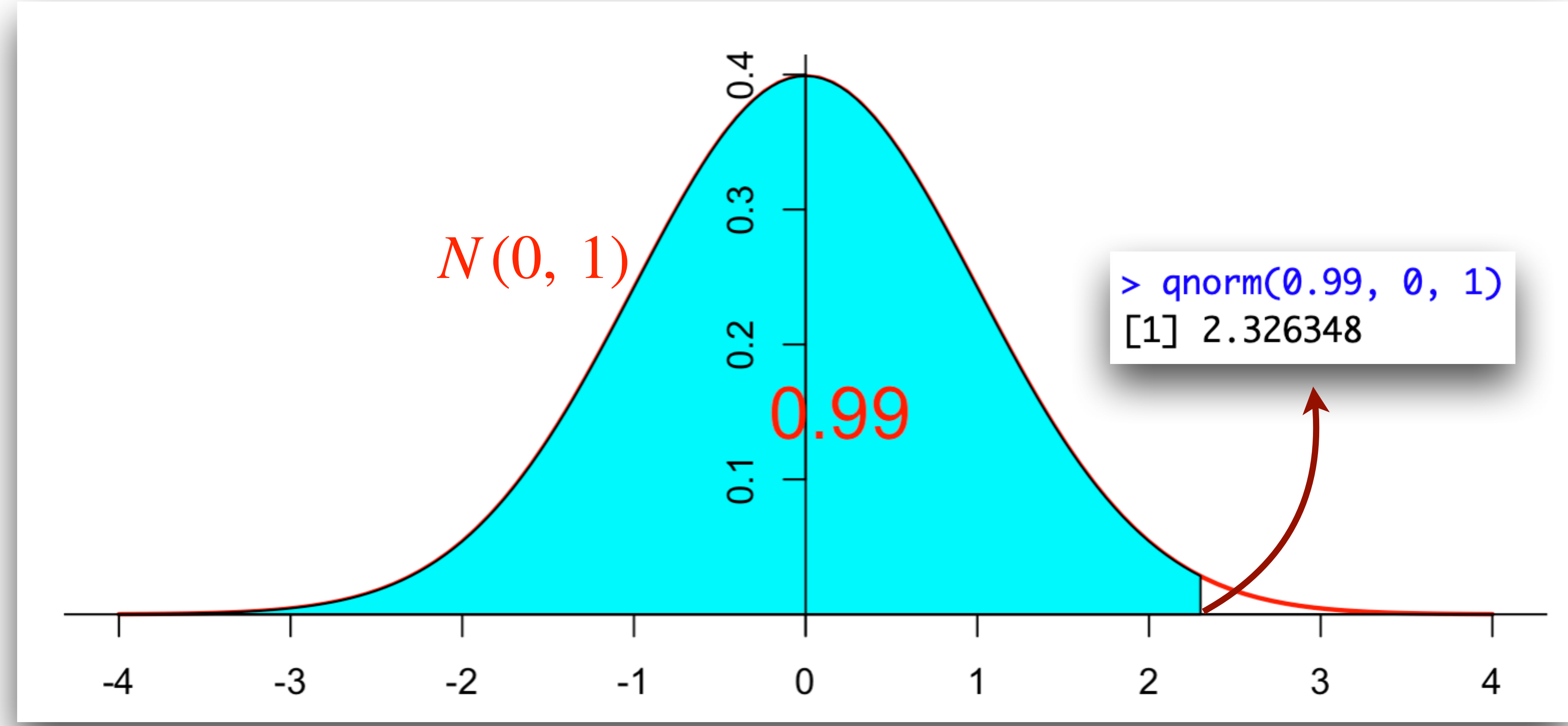
- 常见的连续型随机变量及其分布

② 正态 (normal) 分布：$\xi \sim N\left(\mu, \sigma^2\right)$. 是概率论中最重要的分布

▸ 例：公共汽车车门的高度是按照成年男子与车门顶碰头的机会在 0.01 以下来设计的.

假设成年男子身高 $X \sim N\left(170, 6^2\right)$ (单位：cm)，问车门高度 h 应确定为多少?

$$P\{X \geq h\} \leq 0.01 \quad \Longleftrightarrow \quad P\{X < h\} \geq 0.99$$



$$P\{X < h\} = \Phi\left(\frac{h-170}{6}\right) \geq 0.99$$

$$\Longrightarrow \quad \frac{h-170}{6} \geq 2.326348$$

$$\Longrightarrow \quad h \geq 183.9581 \approx 184 \text{ cm}$$

四、连续型随机变量

- 常见的连续型随机变量及其分布

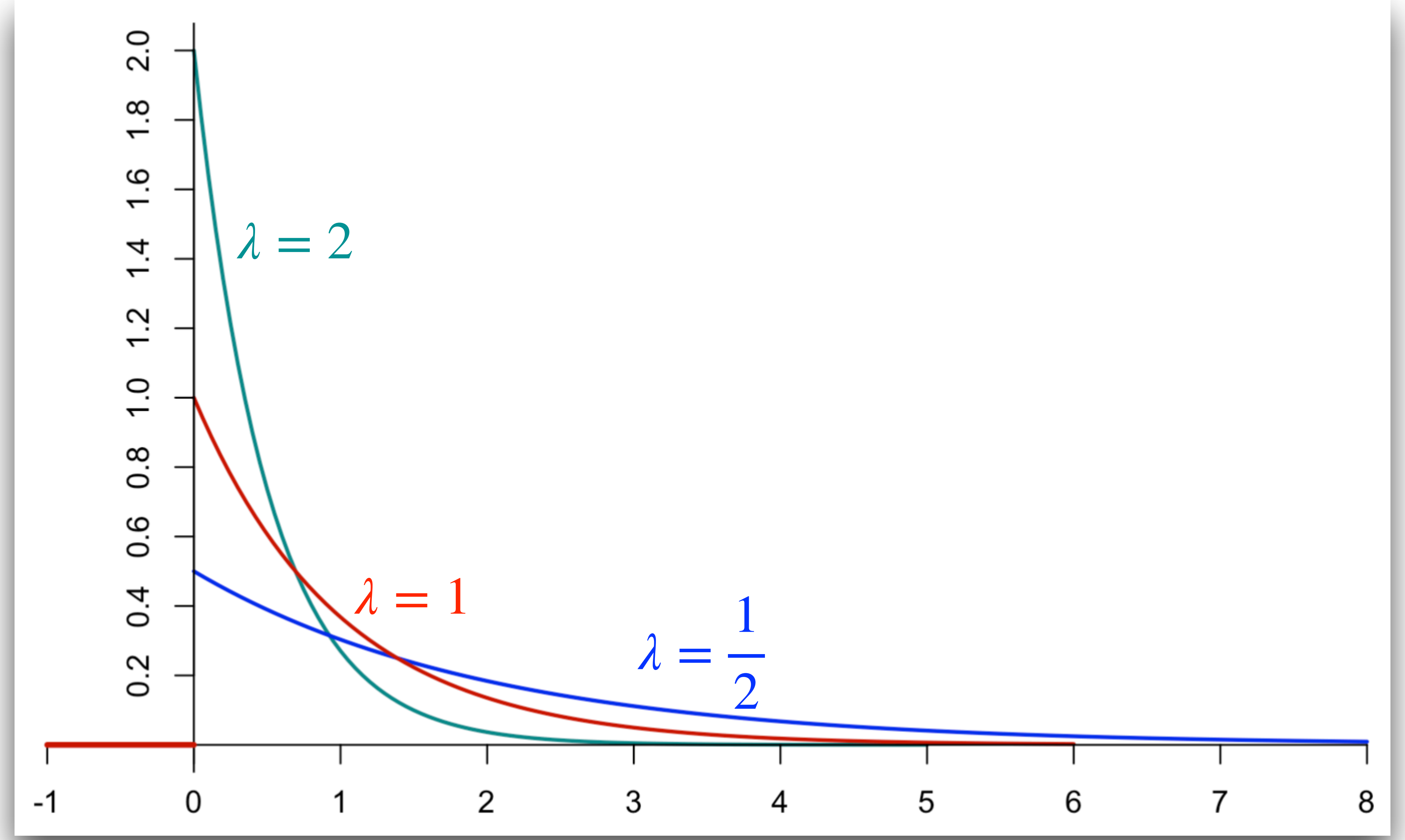
③ 指数 (exponential) 分布：$\xi \sim \mathrm{Exp}\,(\lambda)$.

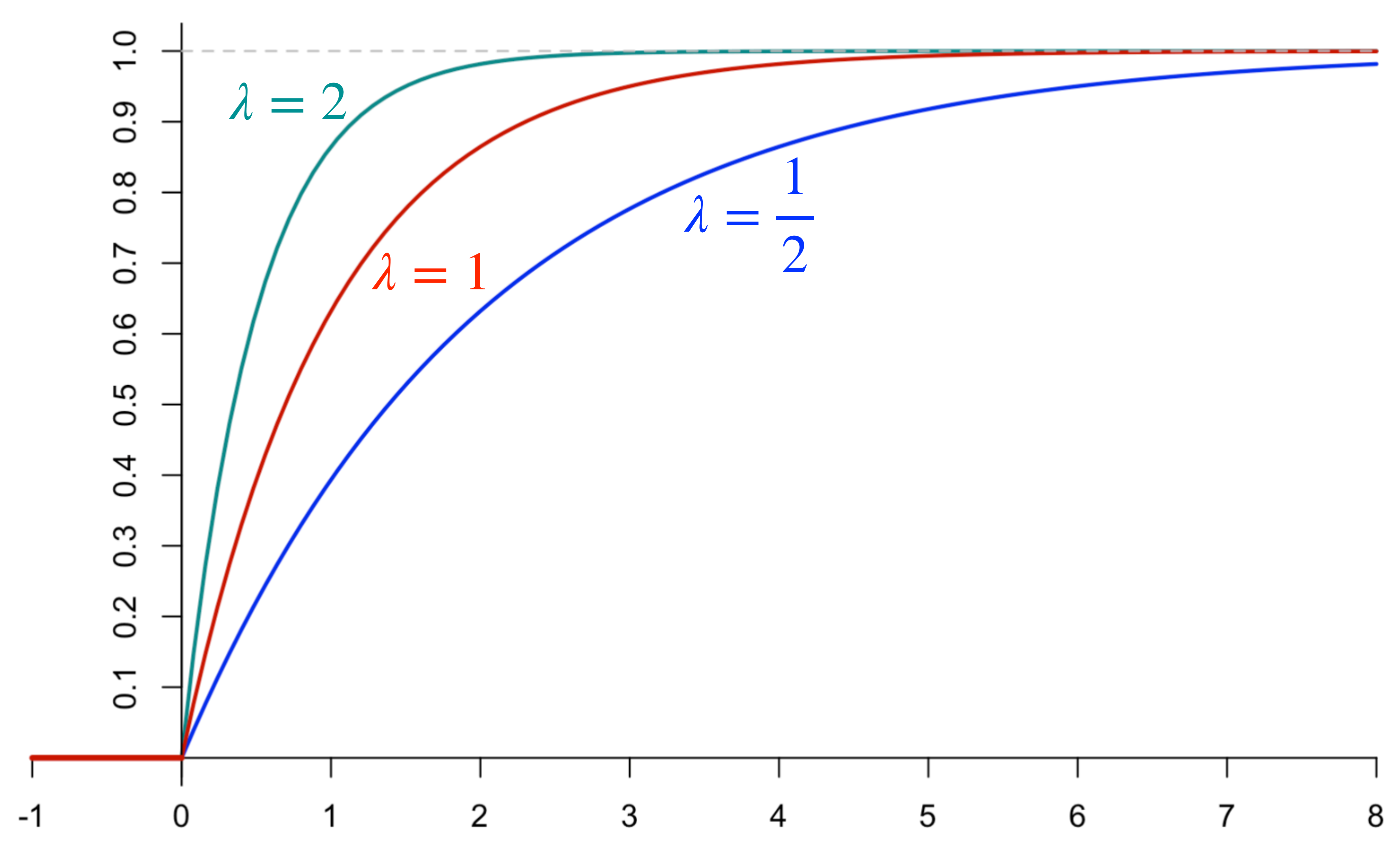
密度函数：$p(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

参数：$\lambda > 0$

分布函数：$F(x) = \int_{-\infty}^{x} p(t)\ \mathrm{d}t$

$$= \begin{cases} 1 - \mathrm{e}^{-\lambda x}, x \geq 0 \\ 0, \quad x < 0 \end{cases}$$





四、连续型随机变量

- 常见的连续型随机变量及其分布

③ 指数 (exponential) 分布：$\xi \sim \text{Exp}(\lambda)$. $p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$; $F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$

▸ 无记忆性：

$$P\left\{\xi\geq s+t \,\middle|\, \xi\geq s\right\}=\frac{P\{\xi\geq s+t\}}{P\{\xi\geq s\}}=\frac{1-P\{\xi<s+t\}}{1-P\{\xi<s\}}=\frac{1-F(s+t)}{1-F(s)}$$

$$=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}=P\{\xi\geq t\}$$

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.

四、连续型随机变量

- 常见的连续型随机变量及其分布

③ 指数 (exponential) 分布：$\xi \sim \text{Exp}(\lambda)$. $p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$; $F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

▸ 定理：取非负实值的随机变量 ξ 服从指数分布，当且仅当它具有无记忆性.

$\Longrightarrow$ 必要性. 我们已证指数分布具有无记忆性.

$\Longleftarrow$ 充分性. 设非负随机变量 ξ 的分布函数为 $F(x)$，记 $G(x) = P\{\xi \geq x\}$

$$\xrightarrow{\text{已知 } \xi \text{ 无记忆}} P\{\xi \geq s+t \mid \xi \geq s\} = P\{\xi \geq t\}$$

$$\xrightarrow{\text{条件概率定义}} = \frac{P\{\xi \geq s+t\}}{P\{\xi \geq s\}}$$

$$0 < a < 1$$

$$\exists\, \lambda > 0 :\ a = e^{-\lambda}$$

$$P\{\xi \geq s+t\} = P\{\xi \geq s\}\cdot P\{\xi \geq t\} \iff G(s+t) = G(s)\cdot G(t) \Longrightarrow G(x) = a^x,\ x \geq 0$$

$$\xi \sim \text{Exp}(\lambda) \quad F(x) = 1 - G(x) = 1 - e^{-\lambda x},\ x \geq 0 \quad \Longrightarrow \quad G(x) = e^{-\lambda x},\ x \geq 0$$

四、连续型随机变量

- 常见的连续型随机变量及其分布

③ 指数 (exponential) 分布：$\xi \sim \text{Exp}\,(\lambda)$. $p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$； $F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$

- 常用于各种寿命分布的近似，如：产品寿命、动物寿命、服务等待时间等.
- 无记忆性表明，无论它已经被使用了多长一段时间，只要还没有损坏，它能再使用一段时间 t 的概率与一件新产品能使用到时间 t 的概率一样，即这种产品将永远年轻.
- 这一点正好说明以指数分布作为寿命分布是有缺陷的. 尽管如此，在有些场合人们还是愿意采用这种易于计算的分布作为产品使用寿命的模型.
- 实际应用当中也可以选正态分布或威利分布作为寿命分布.

四、连续型随机变量

- 常见的连续型随机变量及其分布

③ 指数 (exponential) 分布：$\xi \sim \text{Exp}(\lambda)$. $p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$； $F(x)=\begin{cases}1-e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$

▸ 例：某人乘公交车去上班，等候公交车的时间 (分钟) $\xi \sim \text{Exp}(0.2)$，如果等车时间超过 10 分钟他就乘出租车去上班. 假设一个月工作日为 20 天，求他至少有 3 天是乘出租车上班的概率.

工作日等候时间超过 10 分钟的概率 $p = P\{\xi > 10\} = 1 - P\{\xi \le 10\} = 1 - F(10) = e^{-2}$

$\eta \triangleq \{$每个月等候时间超过 10 分钟的次数$\} = \{$每个月乘出租车上班的次数$\} \Longrightarrow \eta \sim B\left(20, e^{-2}\right)$

至少有 3 天乘出租车上班的概率 $P\{\eta \ge 3\} = \sum_{k=3}^{20}\binom{20}{k}p^k(1-p)^{20-k} = 1-\sum_{k=0}^{2}\binom{20}{k}p^k(1-p)^{20-k}$

$= 0.5206055$

```
> 1 - dbinom(0, 20, exp(-2)) - dbinom(1, 20, exp(-2)) - dbinom(2, 20, exp(-2))
[1] 0.5206055
```

四、连续型随机变量

- 常见的连续型随机变量及其分布

③ 指数 (exponential) 分布：$\xi \sim \text{Exp}(\lambda)$. $p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$; $F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

▸ 指数分布与泊松过程的关系

$\xi(t)$ 表示参数为 λt 的 Poisson 过程

平稳性：$[t_0, t_0 + t)$ 内仅与 t 有关
独立增量性：与 t_0 之前的事件无关
普通性：充分小间隔内至多发生一个

$$P\{\xi(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots$$

$\tau_1 \triangleq \{\xi(t)$ 的第一个跳跃发生的时刻$\}$

$$P\{\tau_1 \geq t\} = P\{\xi(t) = 0\} = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}$$

$$\Longrightarrow \quad P\{\tau_1 < t\} = 1 - P\{\tau_1 \geq t\} = 1 - e^{-\lambda t} \quad \Longrightarrow \quad \tau_1 \sim \text{Exp}(\lambda)$$

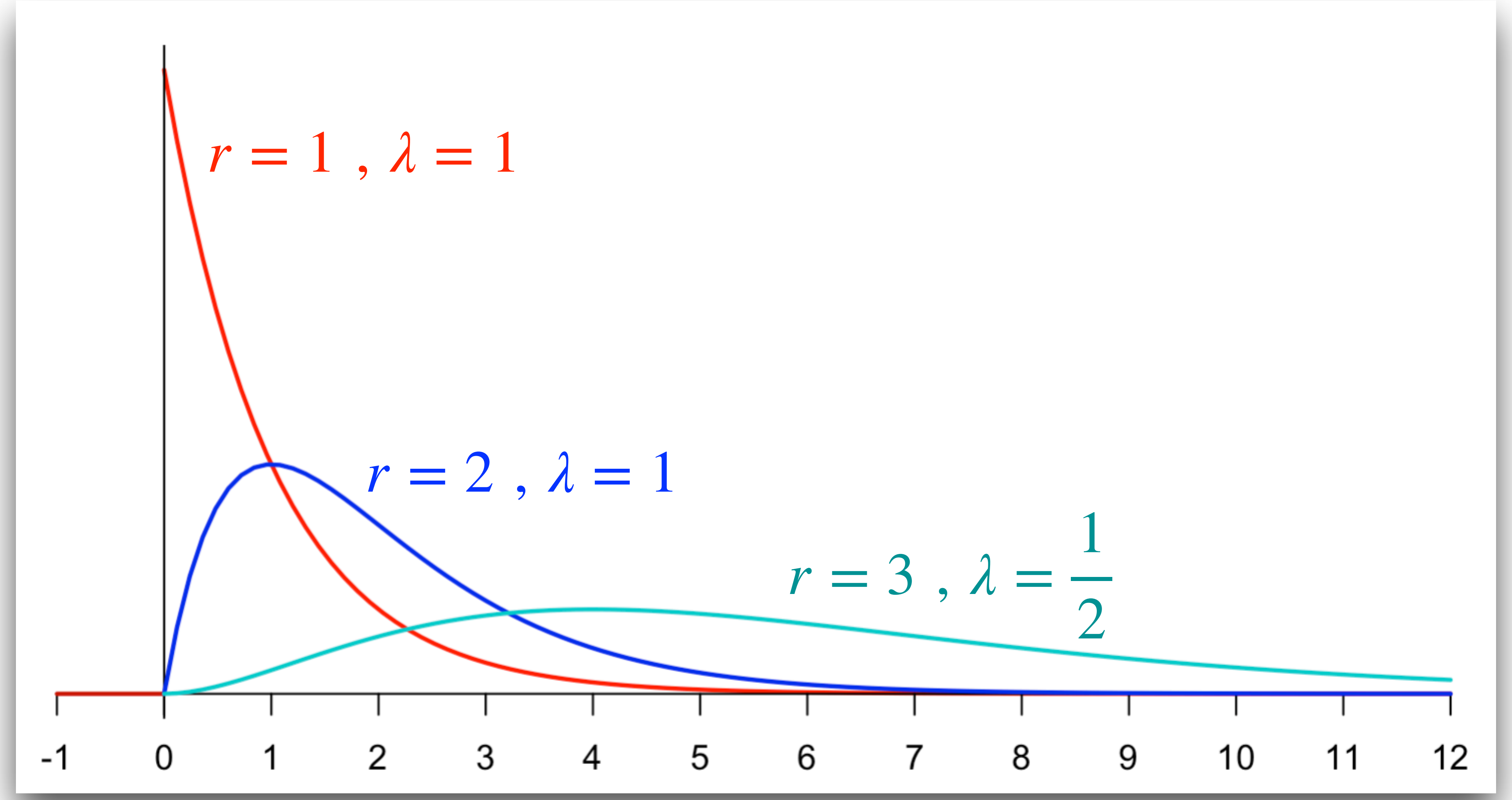
四、连续型随机变量

- 常见的连续型随机变量及其分布

④ Erlang (埃尔朗) 分布：

密度函数：$p(x) = \begin{cases} \dfrac{\lambda^r}{(r-1)!} x^{r-1} \mathrm{e}^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

参数：r 为正整数，$\lambda > 0$ 为实数



四、连续型随机变量

- 常见的连续型随机变量及其分布

④ Erlang (埃尔朗) 分布： $p(x)=\dfrac{\lambda^r}{(r-1)!}x^{r-1}\mathrm{e}^{-\lambda x}\,,\quad x\geq 0\,,\quad \left(r\in\mathbb{N}^+\,,\lambda\in\mathbb{R}^+\right)$

▸ $\xi(t)$ 表示参数为 λt 的 Poisson 过程,

W_r 的分布函数

$$F(t)=P\left\{W_r<t\right\}=P\left\{\xi(t)\geq r\right\}=1-P\left\{\xi(t)<r\right\}=1-\sum_{k=0}^{r-1}\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}$$

$$\Longrightarrow\quad p(t)=\frac{\mathrm{d}}{\mathrm{d}t}F(t)=\sum_{k=0}^{r-1}\frac{\mathrm{e}^{-\lambda t}}{k!}\left[\lambda(\lambda t)^k-k\lambda(\lambda t)^{k-1}\right]=\lambda\mathrm{e}^{-\lambda t}\left[\sum_{k=0}^{r-1}\frac{(\lambda t)^k}{k!}-\sum_{k=1}^{r-1}\frac{(\lambda t)^{k-1}}{(k-1)!}\right]$$

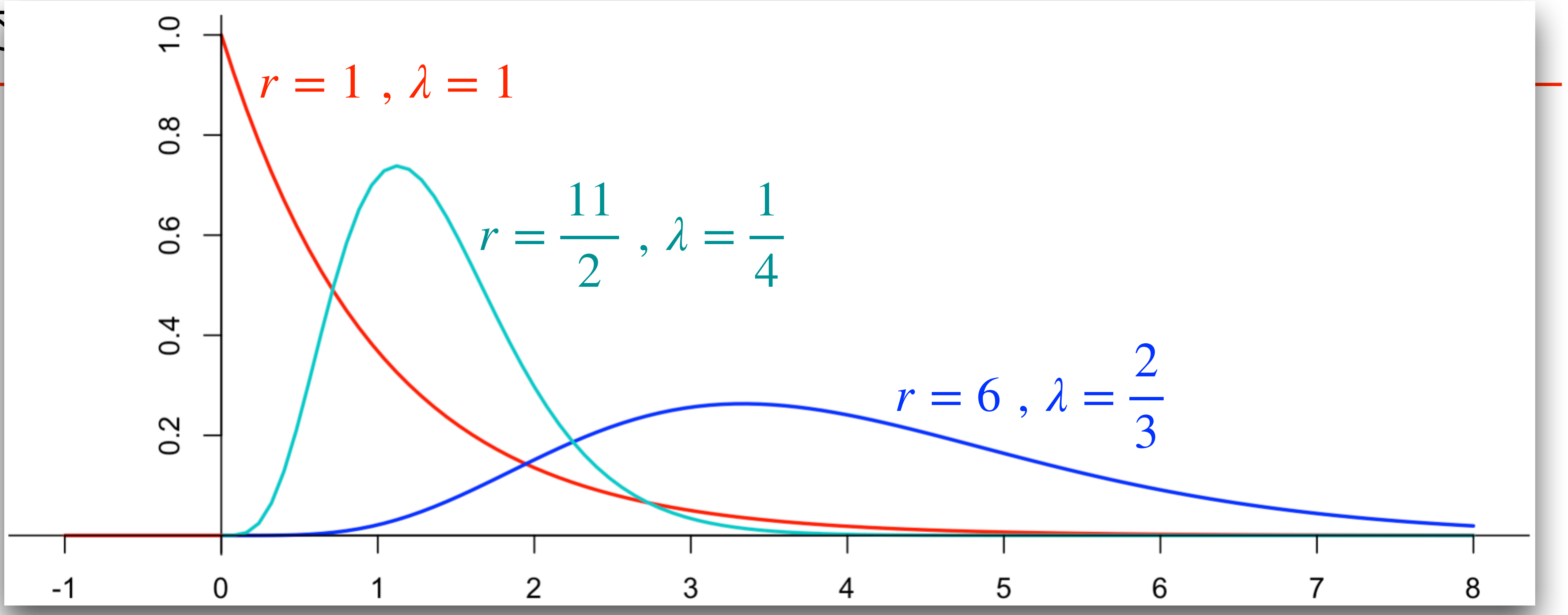
$$=\lambda\mathrm{e}^{-\lambda t}\frac{(\lambda t)^{r-1}}{(r-1)!}=\frac{\lambda^r}{(r-1)!}\,t^{r-1}\,\mathrm{e}^{-\lambda t}$$

Erlang 分布

四、连续型随机变量

- 常见的连续型随机变量及其分布

⑤ Γ (伽马) 分布：$\xi \sim \Gamma(r, \lambda)$.



密度函数：$p(x) = \begin{cases} \dfrac{\lambda^r}{\Gamma(r)} x^{r-1} \mathrm{e}^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$

参数：$r \in \mathbb{R}^+$, $\lambda \in \mathbb{R}^+$

Γ 函数：$\Gamma(r) = \displaystyle\int_0^{+\infty} x^{r-1}\,\mathrm{e}^{-x}\,\mathrm{d}x$, $r \in \mathbb{R}^+$

Γ 函数的性质：

$$\begin{cases} (1) & \Gamma(1) = 1 \\ (2) & \Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi} \\ (3) & \forall\ r \in \mathbb{R}^+,\ \ \Gamma(r+1) = r\,\Gamma(r) \\ (4) & \forall\ r \in \mathbb{N}^+,\ \ \Gamma(r+1) = r! \\ (5) & \dbinom{n+r-1}{n} = \dfrac{\Gamma(n+r)}{\Gamma(n)\Gamma(r)} \\ (6) & \dfrac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} = \displaystyle\int_0^1 u^{r-1}(1-u)^{s-1}\mathrm{d}u \end{cases}$$

▶ 特别：$\Gamma(1, \lambda) = \text{Exp}(\lambda)$.

当 $r \in \mathbb{N}^+$ 时，$\Gamma(r, \lambda)$ 就是 Erlang 分布.

当 $n \in \mathbb{N}^+$ 时，$\Gamma\left(\dfrac{n}{2}, \dfrac{1}{2}\right)$ 就是 $\chi^2(n)$ 分布.

四、连续型随机变量

- 常见的连续型随机变量及其分布
  - Erlang 分布是丹麦数学家 Erlang 在研究电话问题时引进的，其研究开创了排队论学科.
    - 当 $r = 1$ 时，Erlang 分布即为指数分布.
  - Erlang 分布与 Pascal 分布有着类似的性质.

$$\tau_1 = W_1\ ,\ \tau_r = W_r - W_{r-1}\ ,\quad r = 2, 3, \ \ldots$$

Poisson 过程的第 r 个跳跃间隔

$$\Longrightarrow \quad W_r = \tau_1 + \tau_2 + \cdots + \tau_r$$

- 可以证明，$\tau_1, \tau_2, \ \ldots, \tau_r$ 均服从参数为 λ 的指数分布，并且相互独立.

  - Γ 分布包含 Erlang 分布作为特例，在概率论和数理统计中有许多应用，是重要分布之一.

四、连续型随机变量

- 常见的连续型随机变量及其分布

⑥ Bernoulli 过程与 Poisson 过程

- 若每隔 Δt 进行一次实验，则 Bernoulli 试验也可以看作是一个随时间而变化的过程.
- 在 Bernoulli 试验中，到时刻 $n\Delta t$ 为止，共进行了 n 次试验. 成功次数服从二项分布.

  而在泊松过程中，至时刻 t 的来到数则服从泊松分布.
- 为等待第一次成功，Bernoulli 试验中的等待时间服从几何分布（lecture10, p.g.38）；而 Poisson 过程中则服从指数分布（lecture12, p.g. 23）. 它们都具有无记忆性..
- 为等待第 r 次成功，Bernoulli 试验中的等待时间服从 Pascal (负二项) 分布（lecture11, p.g.19）；而 Poisson 过程中则服从 Erlang 分布（lecture12, p.g.25 ）.

3.2 随机向量，随机变量的独立性

一、随机向量及其分布

二、边际 (边缘) 分布

三、条件分布

四、随机变量的独立性

一、随机向量及其分布

- 问题的提出

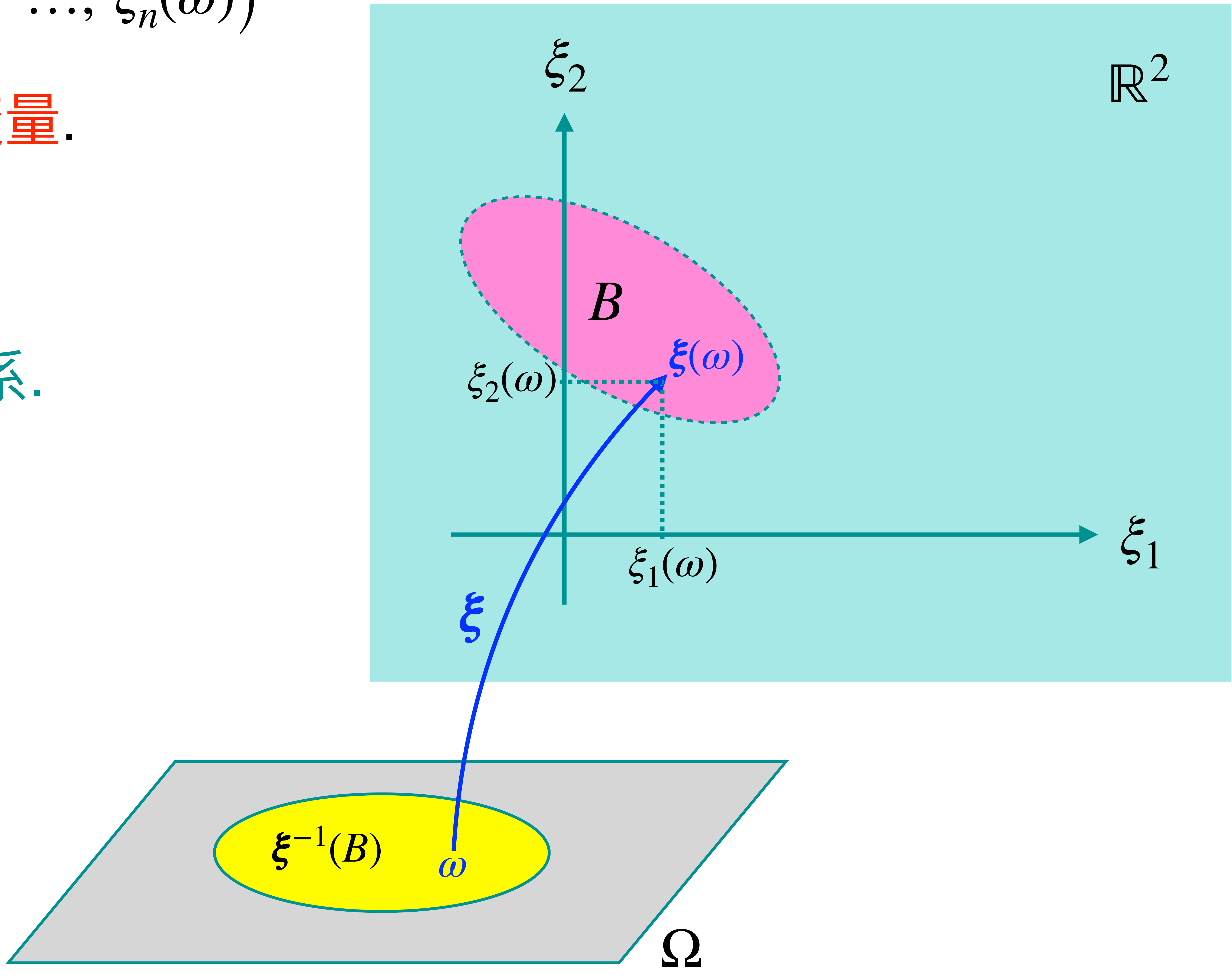
  - 例：研究某一地区学龄儿童的发育情况. 仅研究身高 H 的分布或仅研究体重 W 的分布是不不够的，需要同时考察每个儿童的身高和体重数值，研究身高和体重之间的关系，这就要引入定义在同一样本空间的两个随机变量.
  - 例：研究某种型号炮弹的弹着点分布. 每枚炮弹的弹着点位置需要由横坐标 X 和纵坐标 Y 来确定，而它们是定义在同一样本空间的两个随机变量.

一、随机向量及其分布

- **定义**：若 $\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)$ 是定义在同一概率空间 $(\Omega, \mathscr{F}, P)$ 上的随机变量，则称

$$\boldsymbol{\xi}(\omega) = \big(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)\big)$$

构成一个 n 维随机向量，亦称 n 维随机变量.



- 定义了 n 维随机向量，不仅可以研究每一个随机变量的性质，还可以分析它们之间的联系.

- $\forall\, x_1, x_2, \ldots, x_n \in \mathbb{R}$,

$$\{\xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\}$$

$$= \bigcap_{i=1}^{n} \{\xi_i(\omega) < x_i\} \in \mathscr{F}$$

一、随机向量及其分布

- **定义**：若 $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$ 是定义在同一概率空间 $(\Omega, \mathscr{F}, P)$ 上的随机变量，则称

$$\boldsymbol{\xi}(\omega) = \left(\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)\right)$$

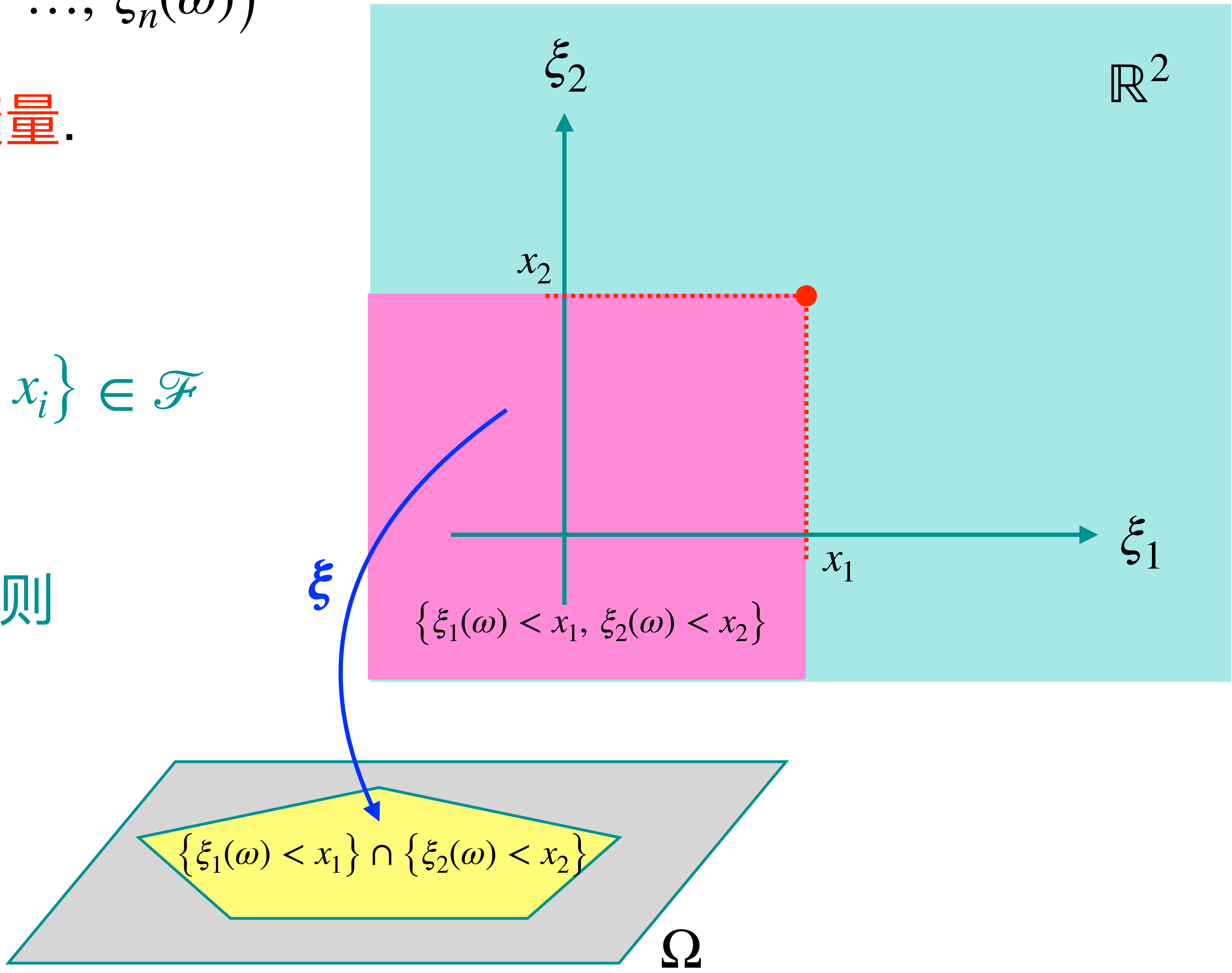
构成一个 n 维随机向量，亦称 n 维随机变量.

- $\forall\, x_1, x_2, \dots, x_n \in \mathbb{R}$,

$$\left\{\xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \dots, \xi_n(\omega) < x_n\right\} = \bigcap_{i=1}^{n} \left\{\xi_i(\omega) < x_i\right\} \in \mathscr{F}$$

- 可以证明：若 B_n 为 $\mathbb{R}^n$ 上任一 Borel 点集，则

$$\left\{\omega \,\middle|\, \boldsymbol{\xi}(\omega) \in B_n\right\} \in \mathscr{F}$$



一、随机向量及其分布

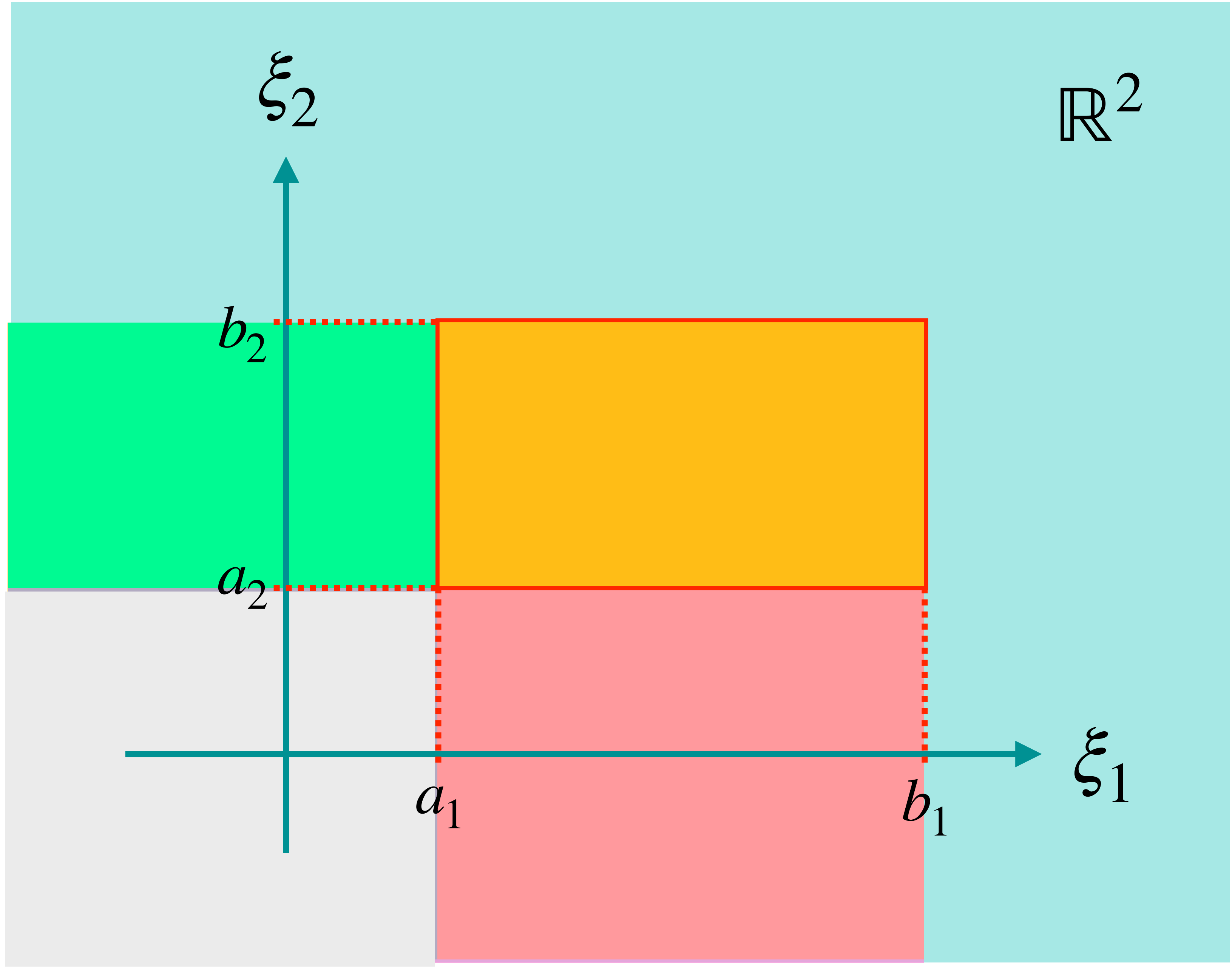
- **定义**：称 n 元函数

$$F\left(x_1, x_2, \ldots, x_n\right) = P\left\{\xi_1(\omega) \le x_1, \xi_2(\omega) \le x_2, \ldots, \xi_n(\omega) \le x_n\right\}$$

为 n 维随机向量 $\boldsymbol{\xi}(\omega) = \left(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)\right)$ 的 (联合)分布函数 (joint cumulative distribution function).

▶ 事件概率的计算：

$$\begin{aligned} P\left\{a_1 < \xi_1 \le b_1, a_2 < \xi_2 \le b_2\right\} = & \ P\left\{\xi_1 \le b_1, \xi_2 \le b_2\right\} \\ & - P\left\{\xi_1 \le a_1, \xi_2 \le b_2\right\} \\ & - P\left\{\xi_1 \le b_1, \xi_2 \le a_2\right\} \\ & + P\left\{\xi_1 \le a_1, \xi_2 \le a_2\right\} \end{aligned}$$



高维亦可计算：$P\left\{a_1 < \xi_1 \le b_1, a_2 < \xi_2 \le b_2, \ldots, a_n < \xi_n \le b_n\right\}$

一、随机向量及其分布

- **多元分布函数的性质**：与一元分布函数类似

④ 成立则 ① 成立，但 ① 成立不能推出 ④ 成立

① 单调性：$F(x_1, x_2, \ldots, x_n)$ 关于每个一元变量 x_i 都是单调不降函数.

② 规范性：
$$\begin{cases} F(-\infty, x_2, x_3, \ldots, x_n) = 0 \\ F(x_1, -\infty, x_3, \ldots, x_n) = 0 \\ \qquad \vdots \qquad\qquad \vdots\ \vdots \\ F(x_1, x_2, \ldots, x_{n-1}, -\infty) = 0 \\ F(+\infty, +\infty, \ldots, +\infty) = 1 \end{cases}$$

后三个性质是二元函数为某二元随机向量的分布函数应满足的条件

③ 右连续性：$F(x_1, x_2, \ldots, x_n)$ 关于每个一元变量 x_i 都是右连续的函数.

④ $n=2$ 时：$\forall\, a_1 \le b_1,\ a_2 \le b_2$ 都有

$$F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \ge 0$$

- **随机向量也有不同类型**：常见的有离散型、连续型.

一、随机向量及其分布

● 常见的多元离散型分布

① 多项 (multinomial) 分布：试验的可能结果为 $A_1, A_2, \ldots, A_r$，记 $P(A_i) = p_i$，$i = 1, 2, \ldots, r$，且

$$p_1 + p_2 + \cdots + p_r = 1$$

独立重复该试验 n 次，定义 $\xi_1, \xi_2, \ldots, \xi_r$ 分别为 $A_1, A_2, \ldots, A_r$ 出现的次数，则

$$P(\xi_1 = k_1, \xi_2 = k_2, \ldots, \xi_r = k_r) = \frac{n!}{k_1!\, k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

$k_i \geq 0$ 取整数，且 $k_1 + k_2 + \cdots + k_r = n$

多项分布用于有放回抽样的场合.

一、随机向量及其分布

- 常见的多元离散型分布

② 多元超几何分布：袋中装有编号为 i 的球 N_i 只，$i=1, 2, \ldots, r$，$N_1+N_2+\cdots+N_r=N$

从中随机取出 n 只球，定义 $\xi_1, \xi_2, \ldots, \xi_r$ 分别表示取出的球中编号为 $1, 2, \ldots, r$ 号球的个数，则

$$P(\xi_1=n_1, \xi_2=n_2, \ldots, \xi_r=n_r)=\frac{\binom{N_1}{n_1}\binom{N_2}{n_2}\cdots\binom{N_r}{n_r}}{\binom{N}{n}}$$

$n_i \geq 0$ 取整数，且 $n_1+n_2+\cdots+n_r=n$

多元超几何分布用于不放回抽样的场合.

一、随机向量及其分布

● 常见的多元连续型分布

▶ 联合概率密度函数 (joint probability density function)： 存在非负函数 $p\left(x_1, x_2, \ldots, x_n\right)$ 使得

$$F\left(x_1, x_2, \ldots, x_n\right) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} p\left(y_1, y_2, \ldots, y_n\right) \mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n$$

▶ 联合概率密度函数满足：

(1) 非负性：$p\left(x_1, x_2, \ldots, x_n\right) \geq 0$

(2) 规范性：$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} p\left(x_1, x_2, \ldots, x_n\right) \mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n = 1$

(1) 与 (2) 构成联合概率密度函数的充分必要条件