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：$P(B)=\sum\limits_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\,\middle|A_i\right)$
    - Bayes 公式：$P\left(A_k\,\middle|B\right)=\dfrac{P\left(A_k\right)\cdot P\left(B\,\middle|A_k\right)}{\sum\limits_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\,\middle|A_i\right)}$

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

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 $(\Omega, \mathscr{F}, P)$，其中

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

    - n 重 Bernoulli 试验 E_n：n 次独立重复的 Bernoulli 试验
    - Bernoulli 分布：$P_1(k) = p^k q^{1-k}\ , \quad k = 0, 1$
    - 二项分布：$b\left(k;\ n, p\right) \triangleq P_n(k) = C_n^k p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}\ , \quad k = 0, 1, 2, \ \ldots, n$
    - 几何分布：$g(k;\ p) = q^{k-1} p = (1-p)^{k-1} p\ , \quad k = 1, 2, \ \ldots$
    - Pascal 分布：$f(k;\ r, p) = C_{k-1}^{r-1} p^r q^{k-r} = \binom{k-1}{r-1} p^r (1-p)^{k-r}\ , \quad k = r, r+1, r+2, \ \ldots$
    - 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 $\begin{cases} P(A_i) = p_i, & i = 1, 2, \ldots, r \\ p_i \geq 0, \text{ 且 } & p_1 + p_2 + \cdots + p_r = 1 \end{cases}$

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

    - 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 $\lfloor (n+1)\cdot p \rfloor$ 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.

已学知识点

- 第二章 条件概率与统计独立性

  - Poisson 分布：$p(k;\ \lambda) = \dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \qquad k = 0, 1, 2, \ldots$

    - 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.

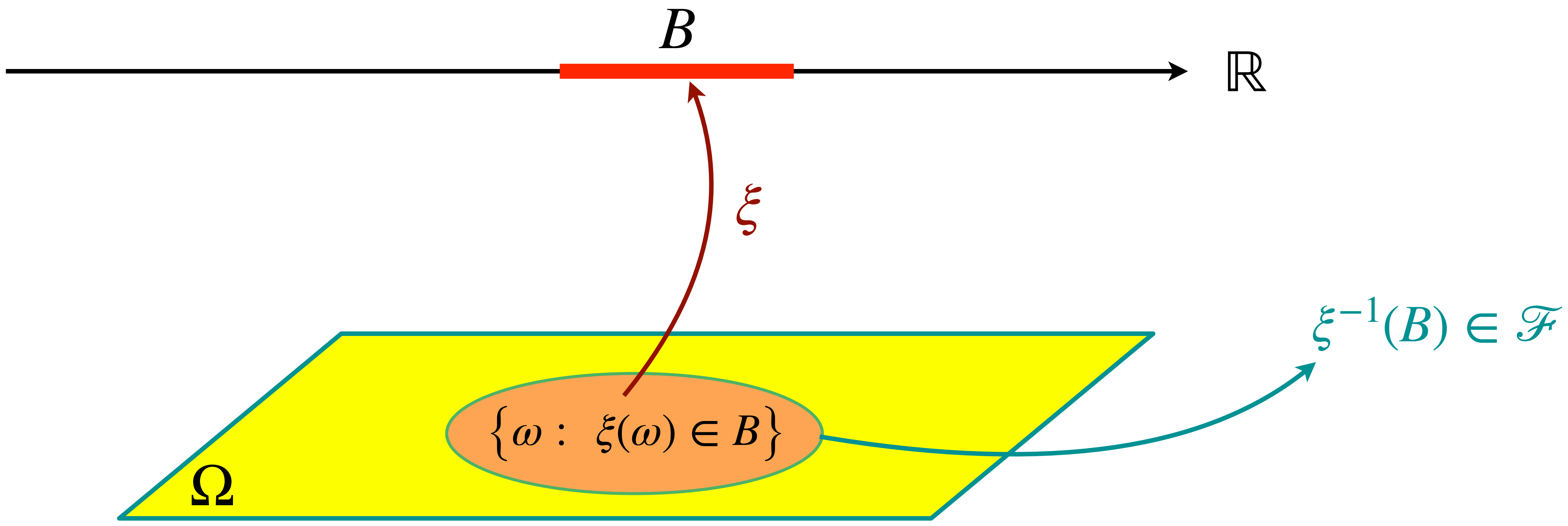
      ③ 普通性：在充分小的时间间隔内，最多发生一次.

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

已学知识点

- 第三章 随机变量与分布函数
  - 利用分布函数计算概率

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

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

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

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

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

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：$P\{\xi = c\} = 1$

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

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

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

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

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

- 几何分布的无记忆性： $P\left\{\eta > m+n \,\middle|\, \eta > m\right\} = P\{\eta > n\} \;, \quad \forall\, m, n \geq 1$
- Remark：几何分布是唯一无记忆的离散型分布.

3.1 随机变量及其分布 (续)

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

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

● 直观解释：

- Bernoulli 试验中，等待首次成功的次数服从几何分布.
- 无记忆性表明：已知试验了 m 次未获得成功，再加做 n 次试验仍不成功的条件概率，等于从开始算起做 n 次试验都不成功的概率. 即，已做过的 m 次失败的试验被忘记了！
- 产生几何分布这种无记忆性的根本原因在于，我们进行的是独立重复试验！
- 如果真的在做独立重复试验，那么不管已经失败过多少次，也不会为今后的试验留下可借鉴的东西.

三、离散型随机变量

- 常见的离散型随机变量及其分布

⑦ Pascal 分布：Bernoulli 试验中，事件 A 第 r 次出现时的试验次数.

$r=1$ 时即为几何分布

$$P\{\zeta = x\} = \binom{x-1}{r-1} p^r q^{x-r}\ , \quad x = r, r+1, \ \ldots$$

- 定义 η_i 表示第 $i-1$ 次成功之后算起，再首次出现成功所需的试验次数，

  则 η_i 服从几何分布 $P\{\eta_i = x\} = q^{x-1}p\ , \quad x = 1, 2, 3, \ \ldots$

- η_i 相互独立，且 $\zeta = \eta_1 + \eta_2 + \cdots + \eta_r$

- 参数为 r 的 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.

三、离散型随机变量

- 常见的离散型随机变量及其分布

$r=1$ 时即为几何分布

⑦ Pascal 分布：Bernoulli 试验中，事件 A 第 r 次出现时的试验次数.

$$P\{\zeta=x\}=\binom{x-1}{r-1}p^r q^{x-r}, \quad x=r, r+1, \ldots$$

$$\binom{-x}{r}=(-1)^r\binom{x+r-1}{r}$$

- 定义 $\tilde{\xi}=\zeta-r$，则 $\tilde{\xi}$ 表示为等待第 r 次成功出现所经历过的失败次数，则

$$P\left\{\tilde{\xi}=y\right\}=P\{\zeta=r+y\}=\binom{r+y-1}{r-1}p^r q^{r+y-r}=\binom{r+y-1}{y}p^r q^y=\binom{-r}{y}p^r(-q)^y, \quad y=0, 1, 2, \ldots$$

- ζ 与 $\tilde{\xi}$ 只是计数方式不同，一个记全部试验数，另一个则只记失败的次数，它们描述的是同样的随机模型，因而本质上是相同的.

不少书以此作为 Pascal 分布的定义

三、离散型随机变量

- 常见的离散型随机变量及其分布

⑦ Pascal 分布：Bernoulli 试验中，事件 A 第 r 次出现时的试验次数.

$r=1$ 时即为几何分布

$$P\{\zeta=x\}=\binom{x-1}{r-1}p^rq^{x-r}, \quad x=r,r+1,\ldots$$

$$\binom{-x}{r}=(-1)^r\binom{x+r-1}{r}$$

▸ 定义 $\tilde{\xi}=\zeta-r$，则 $\tilde{\xi}$ 表示为等待第 r 次成功出现所经历过的失败次数，则

$$P\{\tilde{\xi}=y\}=P\{\xi=r+y\}=\binom{r+y-1}{r-1}p^rq^{r+y-r}=\binom{r+y-1}{y}p^rq^y=\binom{-r}{y}p^r(-q)^y, \quad y=0,1,2,\ldots$$

▸ 特别 $r=1$ 时，考察的是首次出现成功所经历的失败次数 $\tilde{\eta}$，则

也称为几何分布

$$P\{\tilde{\eta}=x\}=q^xp, \quad x=0,1,2,\ldots$$

▸ 有趣的是，等待第 r 次成功所经历的失败次数也可以分解为 $\tilde{\xi}=\tilde{\eta}_1+\tilde{\eta}_2+\cdots+\tilde{\eta}_r$.

三、离散型随机变量

- 常见的离散型随机变量及其分布

⑧ 负二项分布：Pascal 分布中去掉 r 为正整数的限制. $\forall\ r > 0$，称

$$Nb(x;\ r,\ p) = \binom{-r}{x} p^r\,(-q)^x\ ,\quad x = 0,\ 1,\ 2,\ \ldots$$

为负二项 (negative binomial) 分布.

- 负二项分布作为某些离散随机现象的数学模型，已愈来愈引起人们的重视.

四、连续型随机变量

- 定义：设 $F(x)$ 是随机变量 ξ 的分布函数，若存在非负可积函数 $p(x)$，使对任意实数 x 都有

$$F(x)=\int_{-\infty}^{x} p(t)\,\mathrm{d}t$$

则称 ξ 为连续型 (continuous) 随机变量，称 $p(x)$ 为 ξ 的概率密度函数 (probability density function)，简称概率密度或密度函数 (p.d.f.).

- Remark：连续型随机变量的分布函数 $F(x)$ 是绝对连续的，且 $F(x)$ 几乎处处可导，即

$$p(x)=F'(x)=\frac{\mathrm{d}}{\mathrm{d}x}F(x)\,,\qquad \text{for a.e. } x\in\mathbb{R}$$

almost everywhere

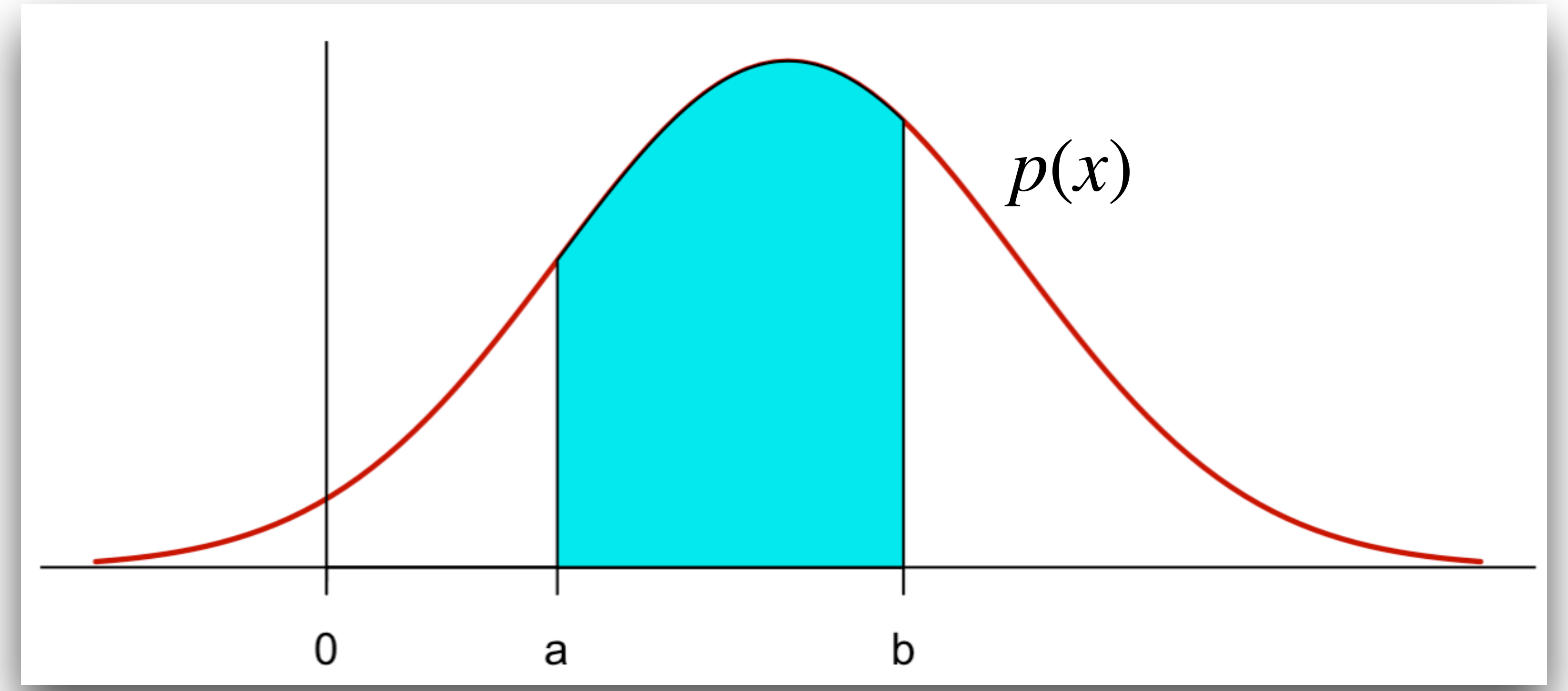
四、连续型随机变量

- 概率密度的性质：

概率密度函数的充分必要条件

① 非负性：$p(x) \geq 0.$

② 规范性：$\displaystyle\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1.$



- 给定密度函数 $p(x)$，$P\{a < \xi \leq b\} = P\{\xi \leq b\} - P\{\xi \leq a\} = F(b) - F(a) = \displaystyle\int_a^b p(x)\,\mathrm{d}x$
- 对 $\mathbb{R}$ 中任一 Borel 点集 B，可以证明 $P\{\xi \in B\} = \displaystyle\int_B p(x)\,\mathrm{d}x\,.$
- 特别，在 $p(x)$ 的连续点 x 处，我们有 $P\{\xi \in (x, x + \Delta x)\} \approx p(x) \cdot \Delta x\,.$
- $\forall\ c \in (-\infty,\ +\infty)$，我们有 $P\{\xi = c\} = 0.$

$$0 \leq P\{\xi = c\} \leq \lim_{h \to 0} \int_c^{c+h} p(x)\,\mathrm{d}x \longrightarrow 0 \qquad (h \to 0)$$

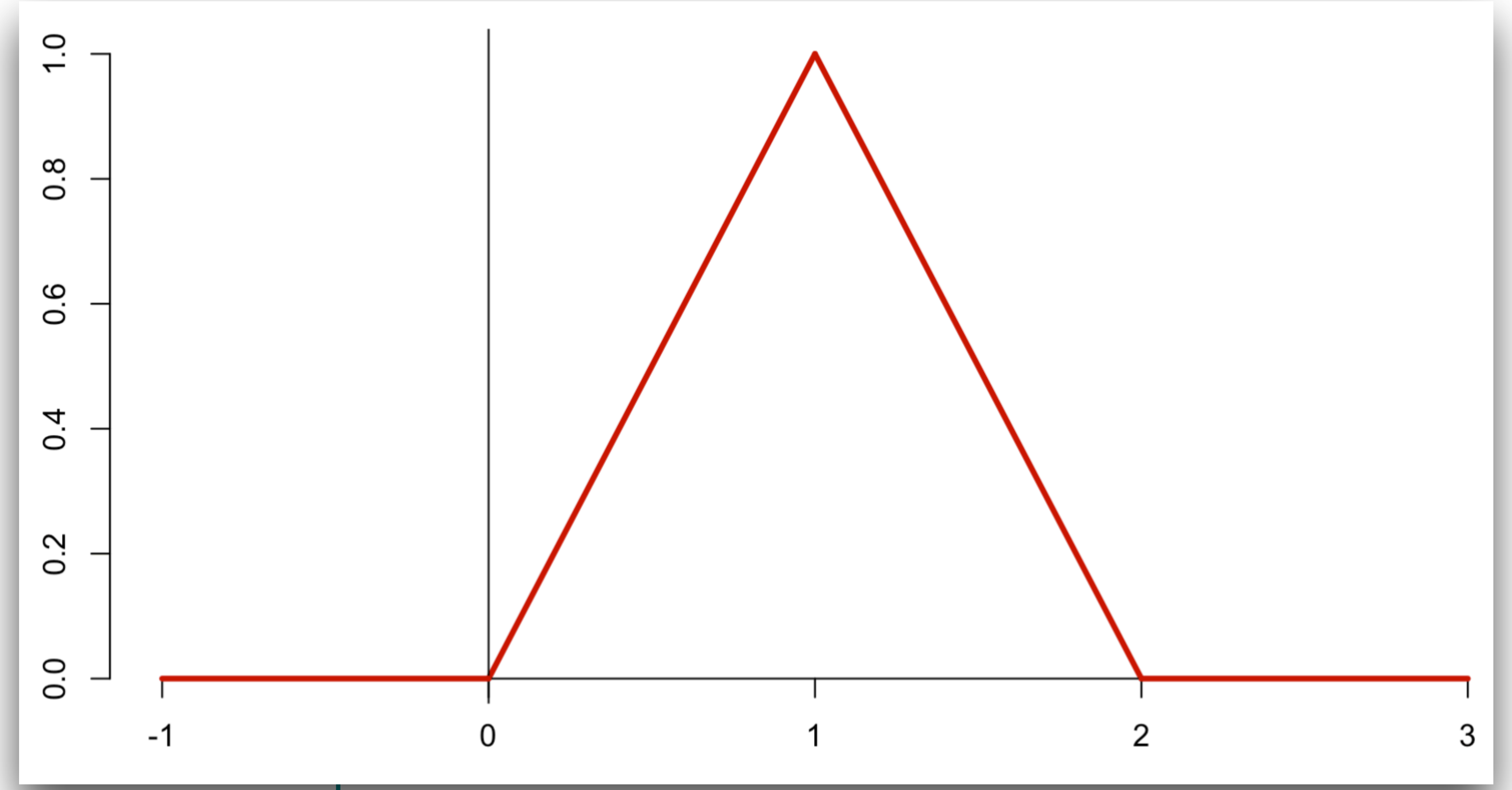
四、连续型随机变量

- **Remark**:
  - 某事件的概率为 0，它并不一定是不可能事件. 某事件的概率为 1，它也不一定是必然事件.
  - 在零测集上可以改变密度函数的值，而不影响分布函数的值.
  - 若两个连续型随机变量的概率密度几乎处处相等，我们即可说二者同分布.
  - 分布函数、概率密度函数，都是对随机变量的概率性质的完整刻划，它们含有相同的信息量，可以相互唯一确定 (几乎处处意义下).
  - 在图形上，密度函数对各种分布特征的显示要优胜得多，因此它比分布函数更为常用.

四、连续型随机变量

- 例: 设 ξ 的 p.d.f. 为

$$p(x) = \begin{cases} ax, & 0 < x < 1 \\ 2 - x, & 1 \le x < 2 \\ 0, & \text{其它} \end{cases}$$



① 求常数 a 的值.

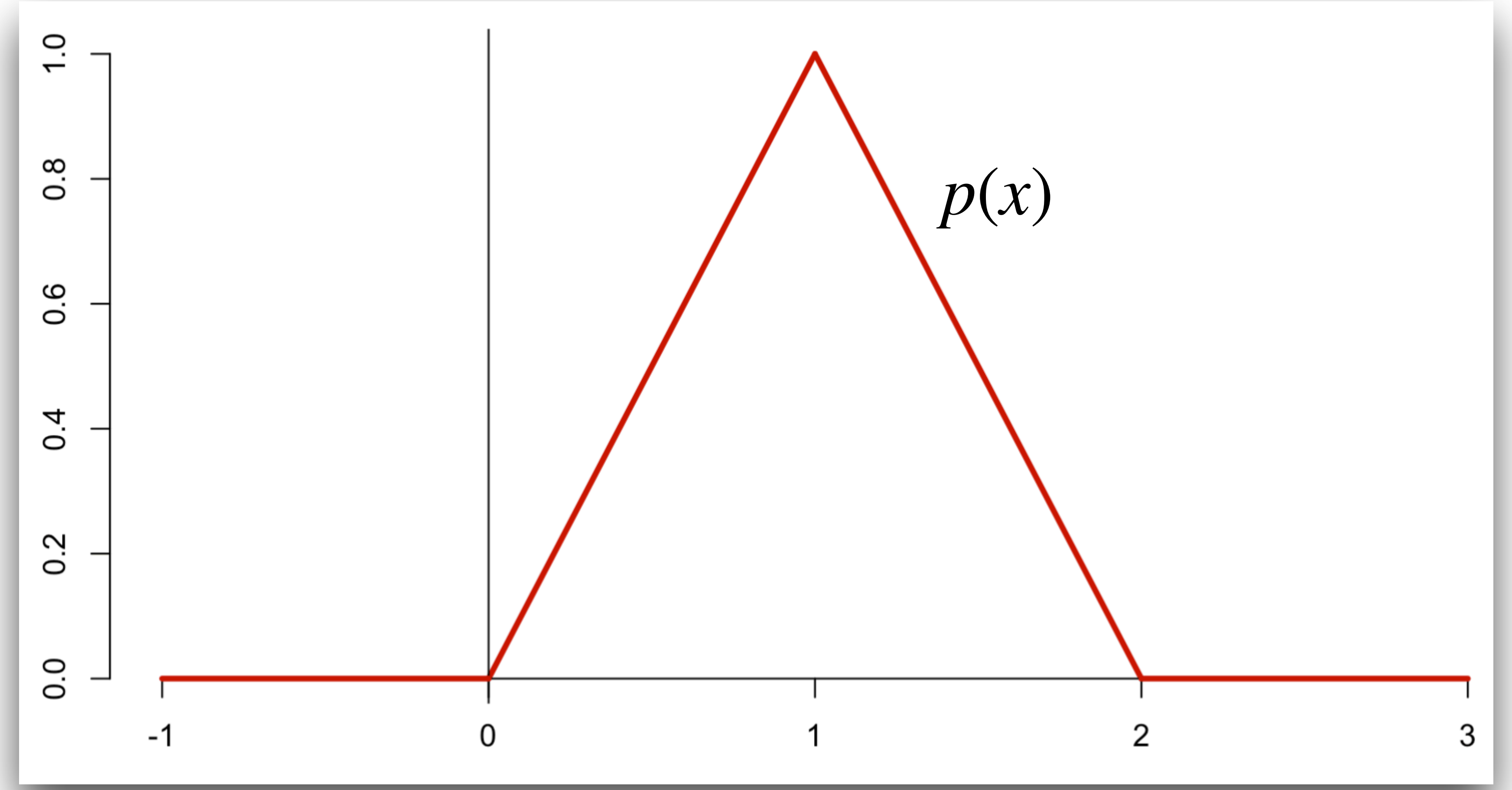
$$1 = \int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = \int_{-\infty}^{0} p(x)\,\mathrm{d}x + \int_{0}^{1} p(x)\,\mathrm{d}x + \int_{1}^{2} p(x)\,\mathrm{d}x + \int_{2}^{+\infty} p(x)\,\mathrm{d}x$$

$$= 0 + \frac{a}{2} + \frac{1}{2} + 0 = \frac{a+1}{2} \quad \Longrightarrow \quad a = 1$$

四、连续型随机变量

- 例: 设 ξ 的 p.d.f. 为

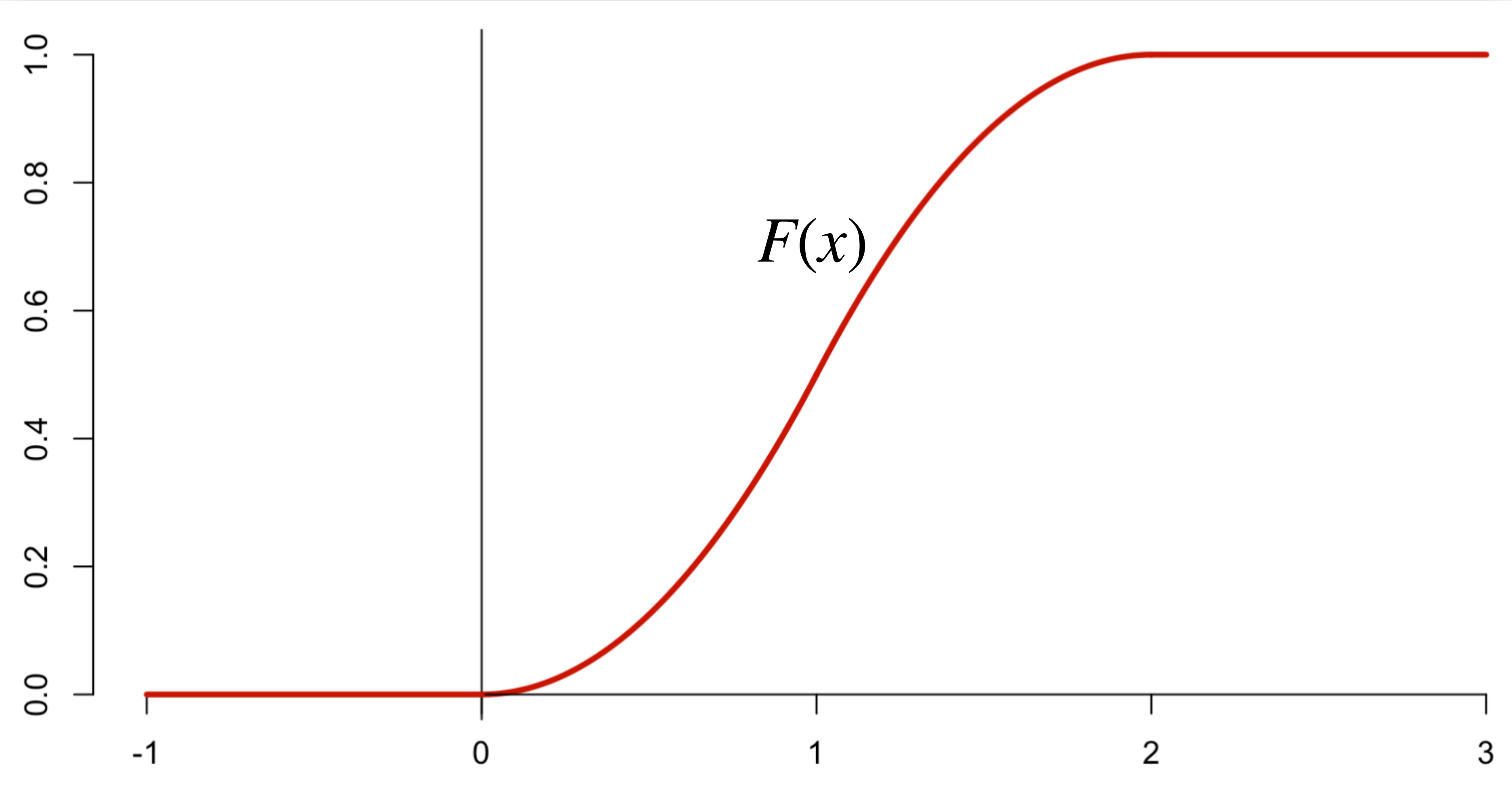
$$p(x)=\begin{cases} x, & 0<x<1 \\ 2-x, & 1\le x<2 \\ 0, & \text{其它} \end{cases}$$



① 求常数 a 的值.

② 写出 ξ 的分布函数. $F(x)=P\{\xi<x\}=\int_{-\infty}^{x} p(x)\,\mathrm{d}x$

$$=\begin{cases} 0, & x\le 0 \\ \int_0^x t\,\mathrm{d}t, & 0\le x<1 \\ \int_0^1 t\,\mathrm{d}t+\int_1^x (2-t)\,\mathrm{d}t, & 1\le x<2 \\ 1, & x\ge 2 \end{cases} = \begin{cases} 0, & x\le 0 \\ \dfrac{x^2}{2}, & 0\le x<1 \\ 2x-\dfrac{x^2}{2}-1, & 1\le x<2 \\ 1, & x\ge 2 \end{cases}$$

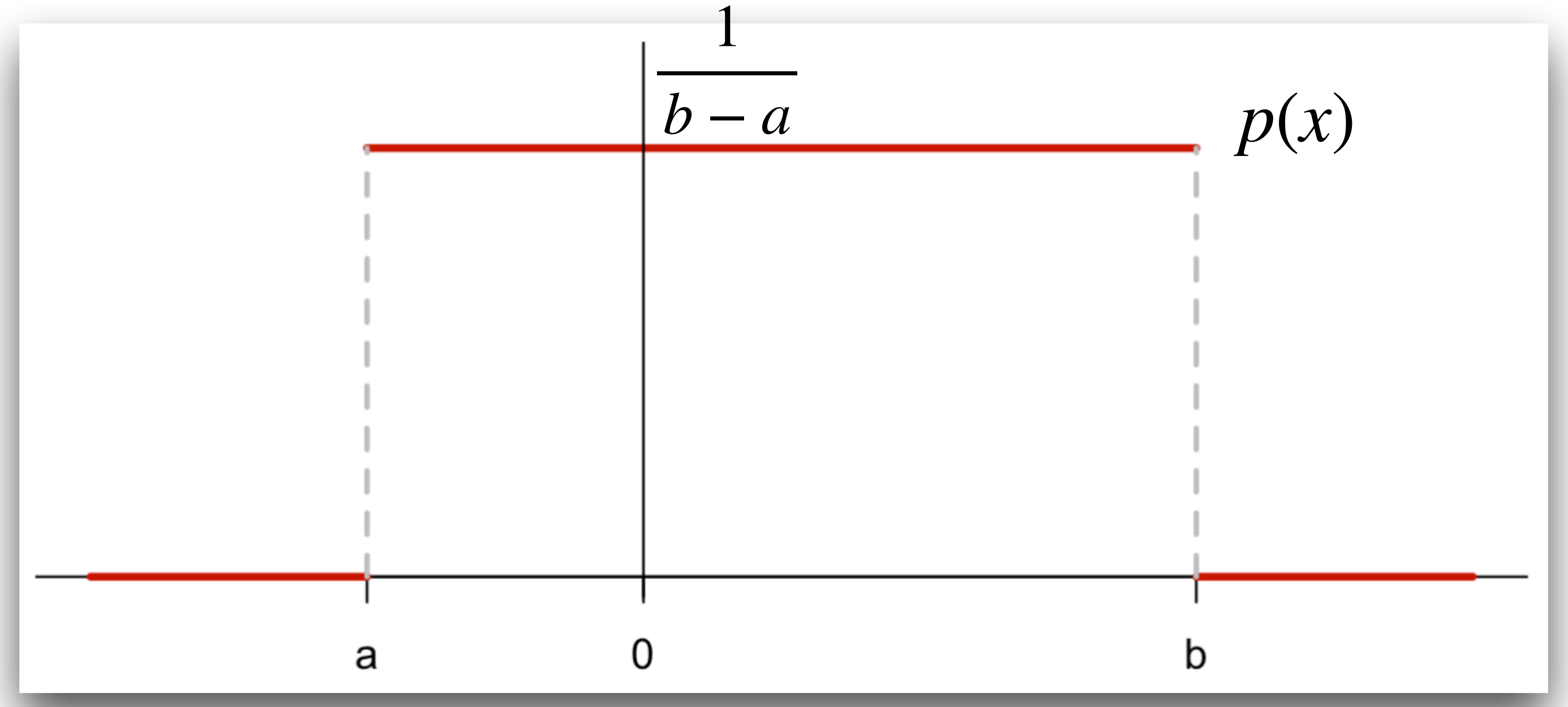


四、连续型随机变量

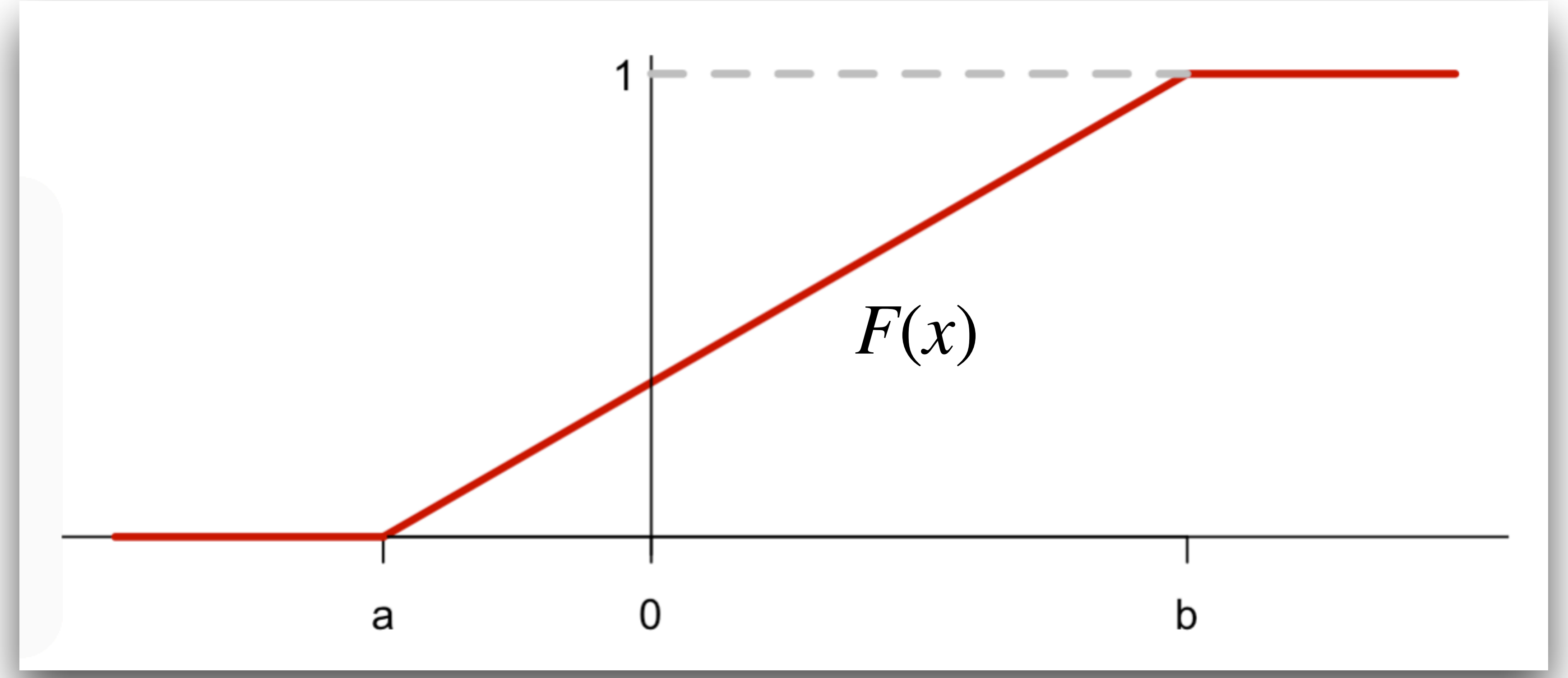
- 常见的连续型随机变量及其分布

① 均匀 (uniform) 分布：$\xi \sim U\left[a, b\right].$

密度函数：$p(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{其它} \end{cases}$



分布函数：$F(x) = \displaystyle\int_{-\infty}^{x} p(t)\,\mathrm{d}t = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}$



四、连续型随机变量

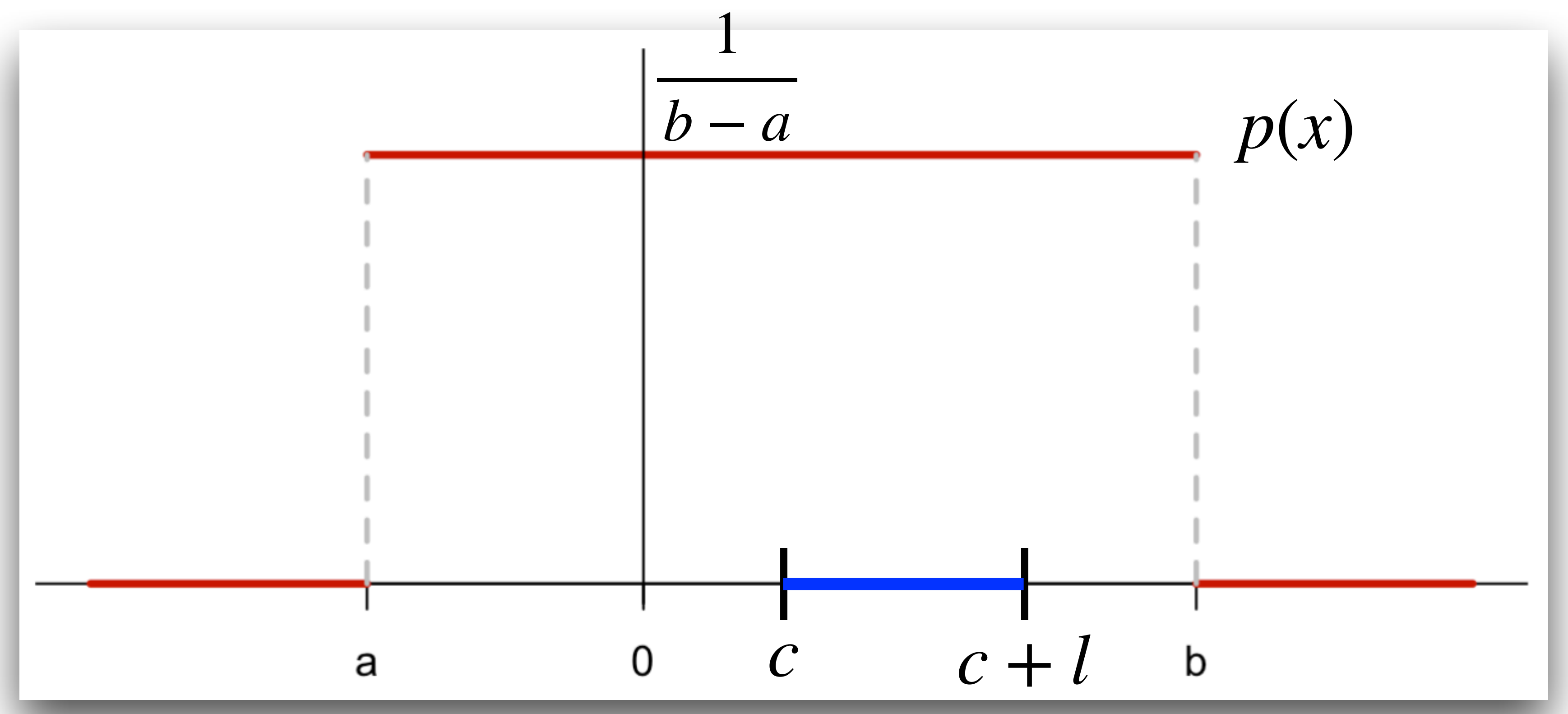
- 常见的连续型随机变量及其分布

① 均匀 (uniform) 分布：$\xi \sim U[a,\, b]$.

▸ 连续情形下的等可能性： $\forall\ [c,\, c+l] \subset [a,\, b]$

$$P\{c \le \xi < c+l\} = \int_c^{c+l} p(x)\,\mathrm{d}x$$

$$= \int_c^{c+l} \frac{1}{b-a}\,\mathrm{d}x = \frac{l}{b-a}$$



▸ 随机变量 ξ 在区间 $[a,\, b]$ 上的任一子区间上取值的概率，与该子区间的长度成正比，而与该子区间的位置无关.

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

正态分布亦称 Gauss 分布

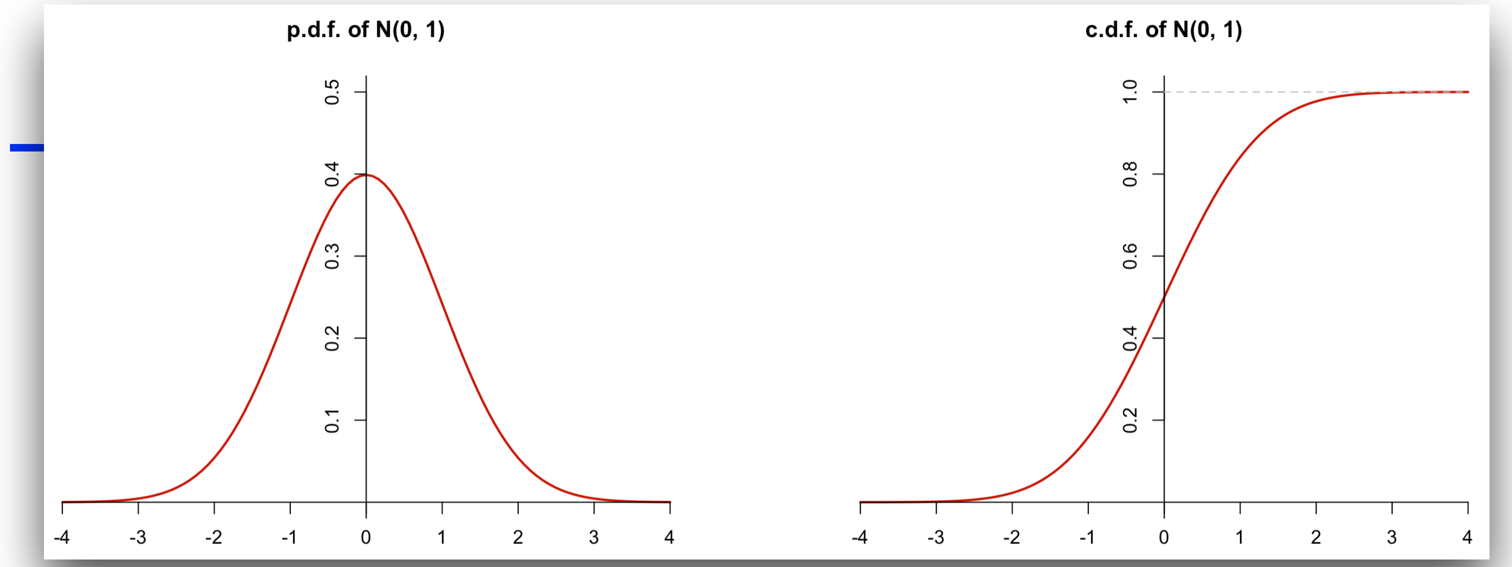
密度函数：$p(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\ \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$ 参数：$-\infty < \mu < +\infty,\ \sigma > 0$

分布函数：$F(x) = \displaystyle\int_{-\infty}^{x} p(t)\ \mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\ \sigma}\ \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\ \mathrm{d}t, \quad -\infty < x < +\infty$

▸ 标准正态分布：$\mu = 0,\ \sigma = 1$，即 $N(0,\ 1).$

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

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



四、连续型随机变量

- 常见的连续型随机变量及其分布

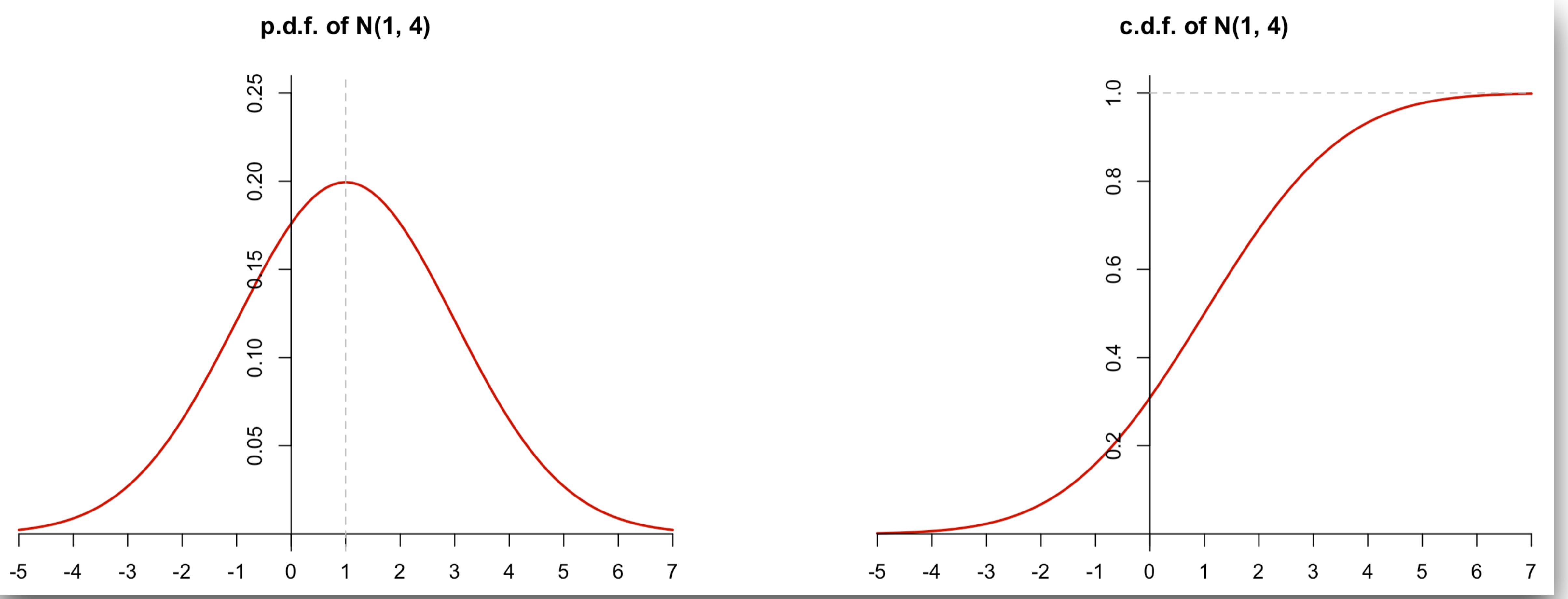
正态分布亦称 Gauss 分布

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

密度函数：$p(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$ 参数：$-\infty < \mu < +\infty\ , \sigma > 0$

分布函数：$F(x) = \displaystyle\int_{-\infty}^{x} p(t)\ \mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\ \sigma}\, \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\, \mathrm{d}t\ , \quad -\infty < x < +\infty$

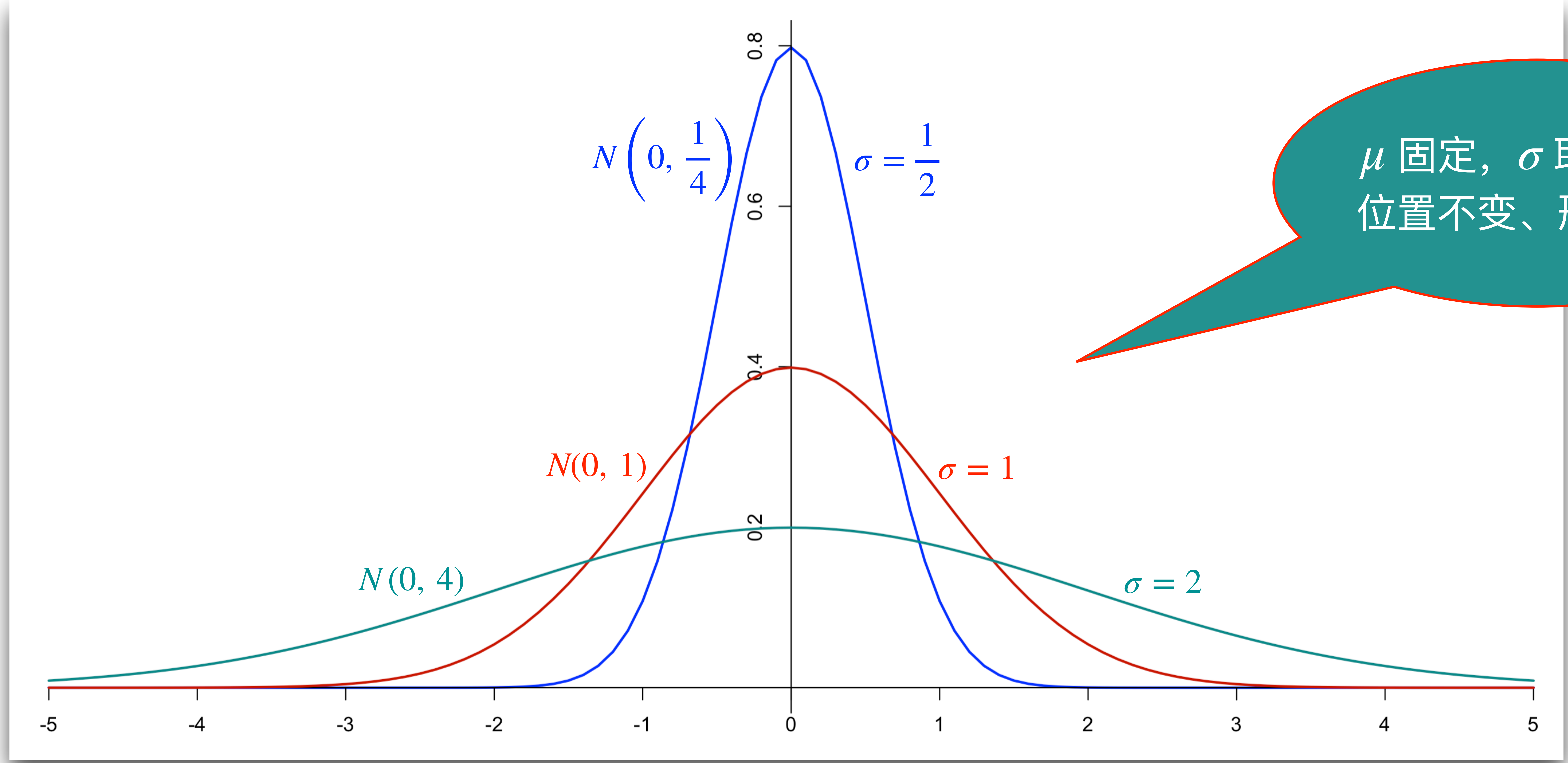
- $p(x)$ 关于 $x = \mu$ 对称.
- $p(x)$ 在 $x = \mu$ 处取最大值 $\dfrac{1}{\sqrt{2\pi}\ \sigma}$.
- μ 决定位置，σ 决定形状.



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：$\xi \sim N\left(\mu,\ \sigma^2\right)$. $p(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\ \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ； $N(0,\ 1)$, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\ \mathrm{e}^{-\frac{x^2}{2}}$

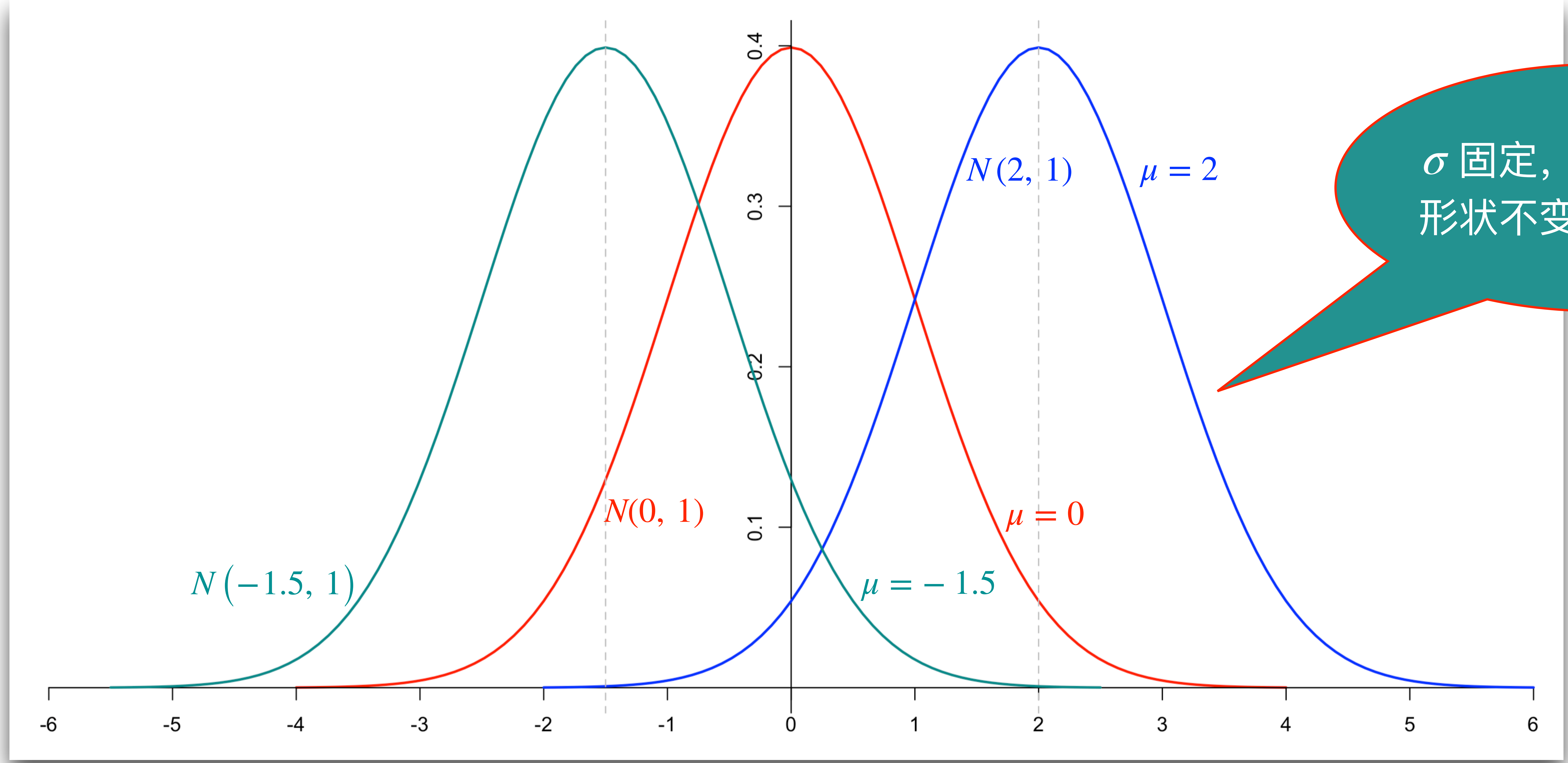


μ 固定，σ 取不同值
位置不变、形状变！

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：$\xi \sim N\left(\mu,\ \sigma^2\right)$. $p(x)=\dfrac{1}{\sqrt{2\pi}\ \sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ； $N(0,\ 1)$, $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$



σ 固定，μ 取不同值
形状不变、位置变!

- 概率密度的性质：
  ① 非负性：$p(x) \geq 0.$
  ② 规范性：$\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1.$

概率密度函数的充分必要条件

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：$\xi \sim N\left(\mu,\, \sigma^2\right).\quad p(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$

▸ 我们来证明 $p(x)$ 定义了一个概率密度函数. 显然有 $p(x) > 0.$

$$z \triangleq \frac{x-\mu}{\sigma}$$

$$\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{z^2}{2}}\,\sigma\,\mathrm{d}z = \frac{1}{\sqrt{2\pi}}\boxed{\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z} = 1$$

$$\left(\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z\right)^2 = \left(\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x\right)\cdot\left(\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{y^2}{2}}\,\mathrm{d}y\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{x^2+y^2}{2}}\,\mathrm{d}x\mathrm{d}y$$

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

$$= \int_{0}^{+\infty}\int_{0}^{2\pi}\mathrm{e}^{-\frac{r^2}{2}}\,r\,\mathrm{d}\theta\,\mathrm{d}r = 2\pi \;\Longrightarrow\; \int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z = \sqrt{2\pi}$$

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：$\xi \sim N\left(\mu, \sigma^2\right)$. $p(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$， $-\infty < x < +\infty$

▸ 可以证明：$\xi \sim N\left(\mu, \sigma^2\right) \Longrightarrow \dfrac{\xi - \mu}{\sigma} \sim N(0, 1)$.

$N(0, 1)$ 的分布函数

$$F(x) = P\left\{\xi < x\right\} = P\left\{\frac{\xi - \mu}{\sigma} < \frac{x - \mu}{\sigma}\right\} = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$P\left\{a < \xi \leq b\right\} = F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P\left\{\left|\xi - \mu\right| < k \cdot \sigma\right\} = P\left\{-k < \frac{\xi - \mu}{\sigma} < k\right\} = \Phi\left(k\right) - \Phi\left(-k\right) = 2\Phi\left(k\right) - 1$$

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：$\xi \sim N\left(\mu,\, \sigma^2\right)$. $p(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < +\infty$

▸ 可以证明：$\xi \sim N\left(\mu,\, \sigma^2\right) \Longrightarrow \dfrac{\xi-\mu}{\sigma} \sim N(0,\, 1)$.

$$\Longrightarrow \begin{cases} P\left\{\left|\xi-\mu\right| < \sigma\right\} = 2\Phi(1) - 1 = 0.6826895 \\ P\left\{\left|\xi-\mu\right| < 2\sigma\right\} = 2\Phi(2) - 1 = 0.9544997 \\ P\left\{\left|\xi-\mu\right| < 3\sigma\right\} = 3\Phi(1) - 1 = 0.9973002 \end{cases}$$

```
2 * pnorm(1, 0, 1) - 1
2 * pnorm(2, 0, 1) - 1
2 * pnorm(3, 0, 1) - 1
```

```
> 2 * pnorm(1, 0, 1) - 1
[1] 0.6826895
> 2 * pnorm(2, 0, 1) - 1
[1] 0.9544997
> 2 * pnorm(3, 0, 1) - 1
[1] 0.9973002
```

$$P\left\{\left|\xi-\mu\right| < k\cdot\sigma\right\} = P\left\{-k < \frac{\xi-\mu}{\sigma} < k\right\} = \Phi\left(k\right) - \Phi\left(-k\right) = 2\Phi\left(k\right) - 1$$

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：$\xi \sim N\left(\mu,\ \sigma^2\right).\quad p(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty < x < +\infty$

▸ 可以证明：$\xi \sim N\left(\mu,\ \sigma^2\right) \Longrightarrow \dfrac{\xi-\mu}{\sigma} \sim N(0,\ 1)\,.$

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

- 正态分布是自然界最常见的一种分布，例如测量的误差，炮弹弹落点的分布，人的生理特征：身高、体重等，农作物的收获量，工厂产品的尺寸：直径、长度、宽度、高度等等都近似服从正态分布.
- 一般来说，若影响某一数量指标的随机因素很多，而每个因素所起的作用都不太大，则这个指标服从正态分布，这一点可以利用概率论的极限定理来证明.
- 正态分布具有许多良好的性质，许多分布可用正态分布来近似，另外一些分布又可以通过正态分布来导出，因此在理论研究中，正态分布发挥着十分重要的作用.

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
```