

第二章 随机变量及分布



连续型随机变量

一. 连续型随机变量的概念与性质

定义 如果对于随机变量 X 的分布函数 $F(x)$, 存在非负函数 $f(x)$, 使得对于任意实数 x , 有

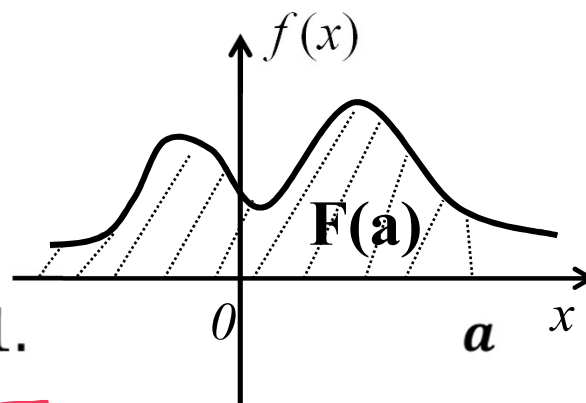
$$F(x) = \int_{-\infty}^x f(t) dt,$$

则称 X 为连续型随机变量. 其中函数 $f(x)$ 称为 X 的概率密度函数, 简称概率密度.

性质

1) 非负性: $f(x) \geq 0$.

2) 归一性: $\int_{-\infty}^{\infty} f(x) dx = 1$.



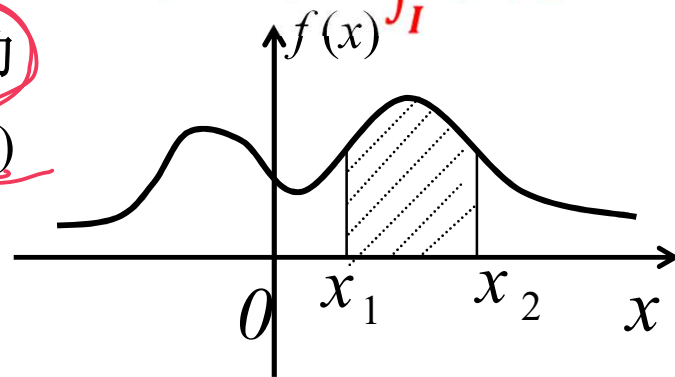
不可能事件 $P=0$ 但 $P=0$ 不一定是不可能事件

反例

3) 连续型随机变量的分布函数是 连续的

若 $f(x)$ 在点 x 处连续, 则有 $F'(x) = f(x)$

$$P(X \in I) = \int_I f(x) dx$$



在点处为0

4) $P(X = a) = F(a) - F(a-) = 0$

$$\begin{aligned} 5) P(x_1 < X \leq x_2) &= F(x_2) - F(x_1) \\ &= \int_{-\infty}^{x_2} f(x) dx - \int_{-\infty}^{x_1} f(x) dx = \int_{x_1}^{x_2} f(x) dx. \\ &= P(x_1 < X < x_2) = P(x_1 \leq X \leq x_2) \\ &= P(x_1 \leq X < x_2) \end{aligned}$$

$$F(x) = \sum_{x_i \leq x} p_i$$

离散

$$= \int_{-\infty}^x f(x) dx.$$

连续

6) $P(x < X \leq x + \Delta x) \approx f(x) \Delta x$ \rightarrow 长条面积

$$P(x < X \leq x + \Delta x) = \int_x^{x+\Delta x} f(t) dt$$

$$= f(\xi) \Delta x, \quad \xi \in (x, x + \Delta x)$$

$$\approx f(x) \Delta x, \quad \Delta x \rightarrow 0.$$

$$p_i \rightarrow f(x) dx$$

$$\frac{p_i}{dx} \rightarrow f(x)$$

$$F(x) = \int_{-\infty}^x f(t) dt$$

例 1 设 X 是连续型随机变量, 其密度函数为

$$f(x) = \begin{cases} c(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{其它} \end{cases}$$

1) X 取值范围(0,2)

$$2) P(X \in I) = \int_I f(x) dx$$

求: (1). 常数 c ; (2). $P(X > 1)$

解: (1). 由密度函数的性质 $\int_{-\infty}^{+\infty} f(x) dx = 1$

$$\text{得 } 1 = \int_{-\infty}^{+\infty} f(x) dx = \int_0^2 c(4x - 2x^2) dx = c \left(2x^2 - \frac{2}{3}x^3 \right) \Big|_0^2 = \frac{8}{3}c$$

$$\text{所以, } c = \frac{3}{8}$$

$$(2) P(X > 1) = \int_1^{+\infty} f(x) dx = \int_1^2 \frac{3}{8}(4x - 2x^2) dx = \frac{1}{2}$$

二、常用的连续型随机变量

1. 均匀分布

2. 指数分布

3. 正态分布

1. 均匀分布

$$X \sim U[a, b]$$

定义 若随机变量 X 的密度函数为

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{其它} \end{cases}$$

则称随机变量 X 服从区间 $[a, b]$ 上的均匀分布. 记作 $X \sim U[a, b]$

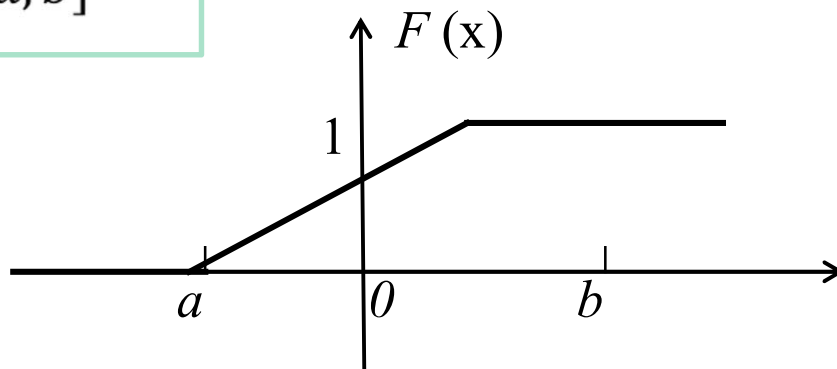
$$P(c < X \leq c+l) = \int_c^{c+l} \frac{1}{b-a} dx = \frac{l}{b-a}$$

a c x $c+l$ b x

向 $[a, b]$ 等可能的投点, 落点 $X \sim U[a, b]$

显然

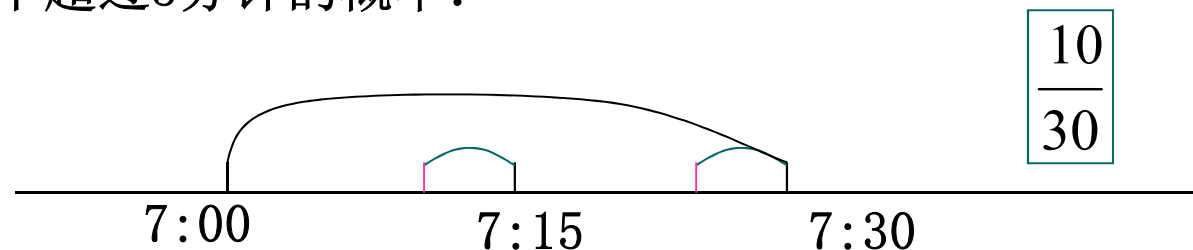
$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & b < x \end{cases}$$



$$F(x) = P(X \leq x) = P(a \leq X \leq x) = \frac{x-a}{b-a}$$

几何概型

例1 设公共汽车站从上午7时起每隔15分钟来一班车, 如果某乘客到达此站的时刻是 7:00 到7:30之间的均匀随机变量. 试求该乘客候车时间不超过5分钟的概率.



$$\frac{10}{30}$$

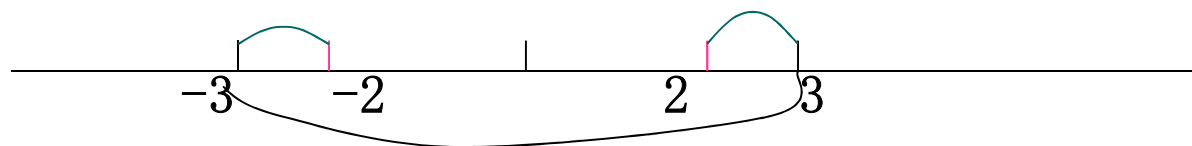
解: 设乘客的到达时刻为 X , 则 $X \sim U[0,30]$

例2 $X \sim U[-3,3]$, 求二次方程 $t^2 - Xt + 1 = 0$ 有实数根的概率

解: 方程有实根的条件是 $\Delta = X^2 - 4 \geq 0$

$$P(X^2 - 4 \geq 0) = P(|X| \geq 2)$$

$$\frac{2}{6}$$

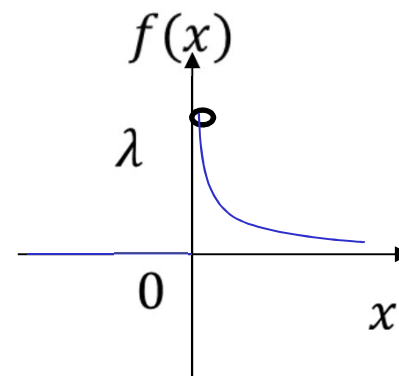


2. 指数分布

$$X \sim e(\lambda) = P(X > x) = e^{-\lambda x}$$

定义: 如果随机变量 X 的密度函数为

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$



其中 $\lambda > 0$ 为常数, 则称随机变量服从参数为 λ 的指数分布.

记为

$$X \sim e(\lambda)$$

密度函数的验证

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^{+\infty} \lambda e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^{+\infty} = 1.$$

指数分布的分布函数

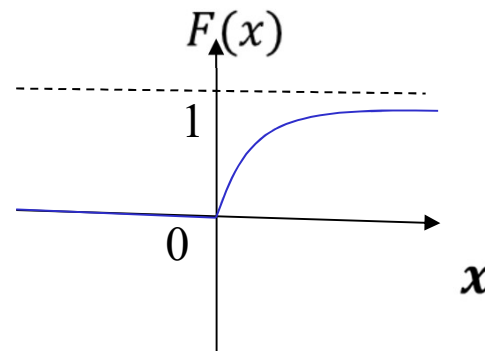
$x \leq 0$ 时,

$$F(x) = 0;$$

$x > 0$ 时,

$$F(x) = \int_0^x \lambda e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^x = 1 - e^{-\lambda x}$$

$$P(X > x) = 1 - F(x) = e^{-\lambda x} \quad (x > 0)$$



指数分布的无记忆性

$$P(X > x) = e^{-\lambda x}$$

若 $X \sim e(\lambda)$ ，则有 $P(X > s+t | X > s) = P(X > t), (s, t > 0)$

$$\text{证明: } P(X > s+t | X > s) = \frac{P(X > t+s, X > s)}{P(X > s)}$$

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

※ **无记忆性是指数分布非常重要的性质**，这条性质决定了指数分布在理论和实践中都有重要的应用，通常用指数分布描述电子元件的寿命和等待时间等指标。

- **指数分布可以看作是等待某个事件首次发生的等待时间。**

- 几何分布的无记忆性

若 $Y \sim G(p)$ ，则 $P(Y > m+n | Y > n) = P(Y > m), (m, n \in \mathbb{Z}^+)$

- **几何分布可以看作是等待某个事件首次发生的等待次数。**

例3 电子元件的寿命服从参数为 $1/100$ 的指数分布（单位：小时）

求 5 个同类型的元件在使用的前 150 小时内恰有 2 个需要更换的概率.

解：已知 $X \sim e(1/100)$.

设 $A = \{\text{元件在使用的前 150 小时内需要更换}\} = \{X \leq 150\}$

设 Y 为 5 个元件中使用寿命不超过150小时的个数.

每次使用一个元件，相当于做了一次独立性试验，使用5个元

件，相当于做了5次独立重复性试验。 $Y \sim B(5, P(A))$

$$\text{则 } P(A) = P\{X \leq 150\} = 1 - e^{-\frac{3}{2}}$$

$$P(Y = 2) = C_5^2 \times \left(1 - e^{-\frac{3}{2}}\right)^2 \times \left(e^{-\frac{3}{2}}\right)^3$$

$$Y \sim B(n, p)$$

独立重复性
试验的次数

A在每次试验
中发生的概率