

# 概率论与数理统计

## 第十一讲 正态分布

## 第11讲 正态分布

### 正态分布是一种重要的分布：

它所揭示的统计规律“两头小，中间大”在自然界与人类社会普遍存在：

- ✓ 测量值与实际值的误差；
- ✓ 分子热运动时每个分子的运动速率；
- ✓ 大气中污染物的浓度；
- ✓ 人群每个个体的智商；
- ✓ .....

# 第11讲 正态分布

## 1. 统计规律的发现

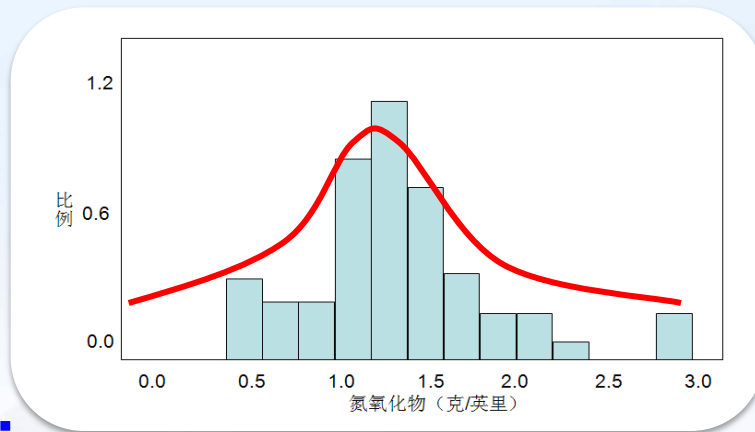
### ● 从数据到直方图

**例（汽车排放物）** 有人研究了46辆车排出的氮氧化物的数量（克/英里）

排放量	0.00-	0.5-	0.75-	1.00-	1.20-	1.30-	1.50-	1.75-	2.00-	2.25-	2.50-
	0.50	0.75	1.00	1.20	1.30	1.50	1.75	2.00	2.25	2.50	3.00
车辆数	3	2	2	9	12	8	3	2	2	1	2

# 第11讲 正态分布

画出所谓直方图如下



**直方图作法:**

- 1.将排放量按范围划分若干小区间 $\Delta_i$ , 设其长度为 $\delta_i$ ;
- 2.统计排放量在此小区间内的汽车频率 $f_i$ ;
- 3.以  $f_i / \delta_i$  为高,  $\delta_i$ 为底作小矩形.

# 第11讲 正态分布

## 2. 正态分布的定义

● **定义** 连续型随机变量  $X$  如果有如下形式的密度函数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\mu \in R, \sigma > 0)$$

则称  $X$  服从参数为  $(\mu, \sigma^2)$  的**正态分布**(normal distribution), 记为  $X \sim N(\mu, \sigma^2)$ .

## 第11讲 正态分布

● **注** 它是否是一个密度函数？

需要验证  $\int_{-\infty}^{\infty} f(x)dx = 1$

事实上，令  $I = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$

$$\text{则 } I^2 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}} dx dy$$

$$\begin{aligned} & \begin{matrix} x=\rho \cos \theta \\ y=\rho \sin \theta \end{matrix} \\ &= \int_0^{2\pi} \int_0^{+\infty} e^{-\frac{\rho^2}{2}} d\rho d\theta = 2\pi \int_0^{+\infty} e^{-\frac{\rho^2}{2}} d\frac{\rho^2}{2} = 2\pi \end{aligned}$$

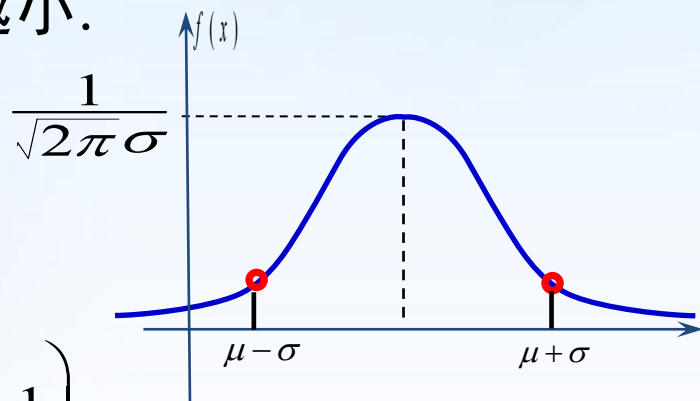
# 第11讲 正态分布

## 3. 密度函数的性质

- ① 关于  $\mu$  对称：在  $x = \mu$  处达到极大值  $\frac{1}{\sqrt{2\pi}\sigma}$ ，越远离  $\mu$ ，密度函数越小。

② 
$$\frac{d^2}{dx^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
$$= \frac{1}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left( \frac{(x-\mu)^2}{\sigma^2} - 1 \right),$$

故  $x = \mu \pm \sigma$  是  $f(x)$  的拐点。



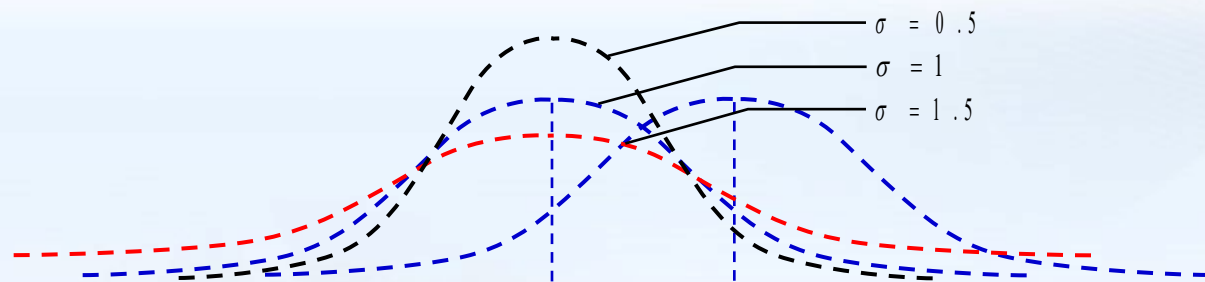
## 第11讲 正态分布

● **注**  $\mu$ 与 $\sigma^2$ 的意义今后会解释.

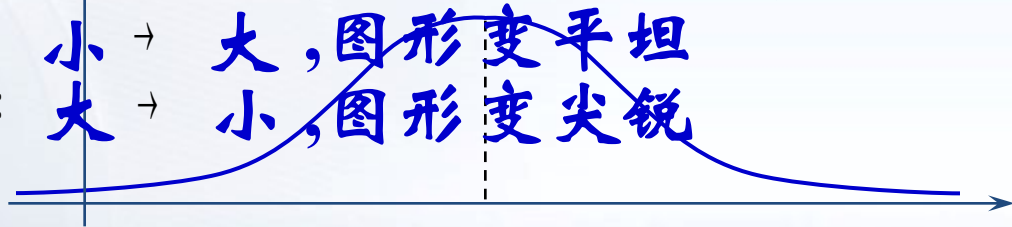
观察 $\mu, \sigma^2$ 的变化对密度函数图形变化的影响.



# 第11讲 正态分布



$\mu$  : 小  $\rightarrow$  大, 图形向右平移, 形状不变  
 $\mu$  : 大  $\xrightarrow{f(x)}$  小, 图形向左平移, 形状不变  
 $\sigma$  : 小  $\rightarrow$  大, 图形变平坦  
 $\sigma$  : 大  $\rightarrow$  小, 图形变尖锐



# 第11讲 正态分布

## 3. 密度函数的性质

**例** (男人身高) 设成年男子身高( $m$ )  $X \sim N(1.70, 0.1^2)$ , 随机观察一男子, 求其身高超过1.80的概率.


易知 
$$P\{X > 1.80\} = \int_{1.80}^{+\infty} \frac{1}{\sqrt{2\pi} \times 0.1} e^{-\frac{(x-1.70)^2}{2 \times 0.1^2}} dx = ?$$

一般地形如  $\int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$  的积分得不到解析表达, 只能作数值积分, 对每个  $x$  制表, 列出结果.

$x$  实际取值: 按适当步长, 离散取  $x_1, x_2, \dots$

## 第11讲 正态分布

若  $\mu, \sigma^2$  变化, 怎么办?

考虑如下变换:  $Z = \frac{X - \mu}{\sigma}$   (称为  $X$  的标准化)

**定理** 设  $X \sim N(\mu, \sigma^2)$ , 则  $Z \sim N(0, 1)$ .

证: 令  $Z$  的分布函数和密度函数分别为  $\Phi(x)$ ,  $\phi(x)$ ,

$$\text{则 } \Phi(x) = P\{Z \leq x\} = P\{X \leq \sigma x + \mu\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\sigma x + \mu} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

$$\text{令 } \frac{t - \mu}{\sigma} = u,$$

$$\text{则 } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du, \quad \text{故 } \phi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

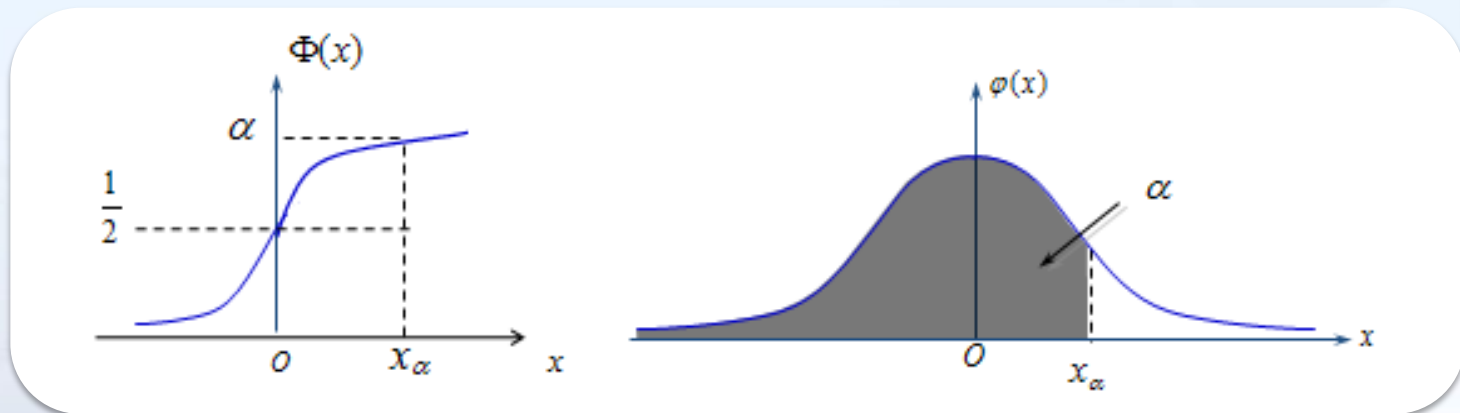
# 第11讲 正态分布

故只需对  $\Phi(x)$  制表, 列出结果.  $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = P\{Z \leq x\}$

$x$	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9983	0.9984	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

## 第11讲 正态分布

● **注** 称 $N(0, 1)$ 为标准正态分布, 其分布函数与密度函数图形如下:



实际问题中, 常有先给出  $\Phi(x) = \alpha$ , 求反函数  $\Phi^{-1}(x) = \alpha$ , 称  $x_\alpha$  为  $\alpha$  **分位数**(quantile).

## 第11讲 正态分布

重作男人身高问题：

$$\begin{aligned} P\{X > 1.80\} &= P\left\{\frac{X - 1.70}{0.1} > \frac{1.80 - 1.70}{0.1}\right\} \\ &= P\{Z > 1\} = 1 - \Phi(1) = 1 - 0.8413 \approx 0.1587. \end{aligned}$$

**例**（男人身高，续）城市公交车门高度设计中, 受结构限制不能太高. 但为了保证95%的男人都能顺利等车, 应保证车门的高度不低于多少？

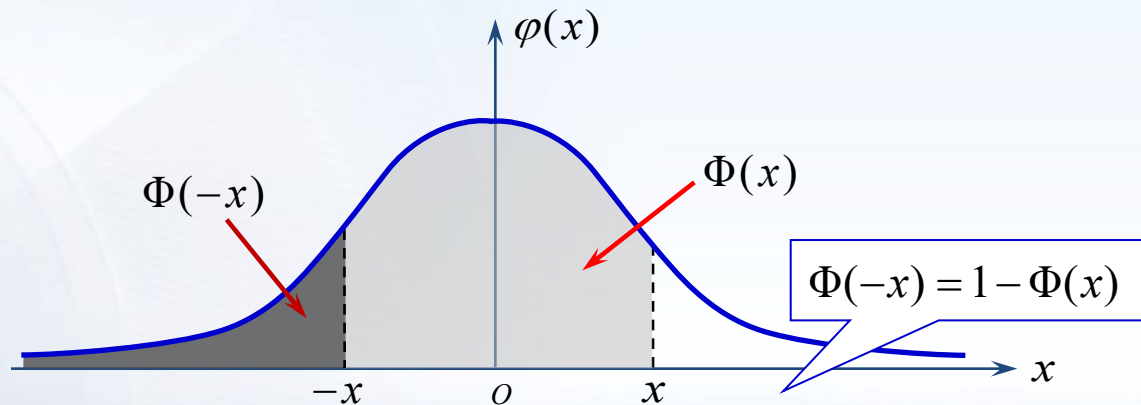
**答:**

$$\begin{aligned} 0.95 &= P\{X < x\} = P\left\{\frac{X - 1.70}{0.1} < \frac{x - 1.70}{0.1}\right\} \\ \Rightarrow \frac{x - 1.70}{0.1} &= 1.65 \quad x = 1.865. \end{aligned}$$

# 第11讲 正态分布

## 问题

1.  $\Phi(x)$ 表中, 通常只列出  $x=3.9$ , 为什么? (细尾)
2.  $\Phi(x)$ 表中, 没有  $x < 0$  的值(对应的  $a$  没有  $a < 0.5$ ), 怎么办? (对称性)  $\Phi(-x) = 1 - \Phi(x)$



## 第11讲 正态分布

### ● 注

工程上采用的  $3\sigma$  法则：认为正态变量的取值在  $\mu \pm 3\sigma$  之间.

设  $X \sim N(\mu, \sigma^2)$ ,

则  $P\{|X - \mu| \leq 3\sigma\} = 0.9974$ .

$3\sigma$  法则常用于产品质量控制.



# 第11讲 正态分布



注

具有类似正态特征的随机变量的密度函数可以很任意, 为什么正态分布的地位最突出.

## 主要考虑

- a) 密度函数有很好的数学性质(无穷可微,  $x^n e^{-x^2}$  可积,  $\forall n$  );
- b) 中心极限定理将揭示的规律. (普适, 和谐!)