3.1. 方差和标准差

- 方差是刻画随机变量在位置参数附近的散布程度.
- 设$EX=\mu$, 偏离量$(X-\mu)$也是随机变量.

$$\begin{cases} |X-\mu|---E|X-\mu|: & \text{平均绝对偏差} \\ (X-\mu)^2---E(X-\mu)^2: & \text{方差} \end{cases}$$

对一般的随机变量, 其方差定义为

定义 4.7 方差和标准差

设随机变量 X 是平方可积的, 即满足 $\mathbb{E}X^2 < \infty$, 则

$$\sigma^2 = \mathrm{Var}(X) \equiv \mathbb{E}(X-\mu)^2, \tag{4.11}$$

$$\sigma = \sqrt{\mathrm{Var}(X)} \tag{4.12}$$

分别称为随机变量 X 的方差和标准差. 也可以称为随机变量分布函数的方差和标准差. ♣

方差的性质

1. $0 \le \mathrm{Var(X)} = \mathrm{EX^2} - \mathrm{(EX)^2}$，故$\mathrm{Var(X)} \le \mathrm{EX^2}$。
2. 常数的方差为0。
3. 若c为常数，则$\mathrm{Var(X+c)} = \mathrm{Var(X)}$。
4. 若c为常数，则$\mathrm{Var(cX)} = \mathrm{c^2Var(X)}$。
5. **独立**随机变量和的方差等于随机变量方差的和，即若$X_1, X_2, \ldots, X_n$相互独立，则 $\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$.
6. $X_1, X_2, \ldots, X_n$相互独立, $c_1, c_2, \ldots, c_n$为n个常数时，有

$$\mathrm{Var}(\sum_{i=1}^n c_i X_i) = \sum_{i=1}^n c_i^2 \mathrm{Var}(X_i).$$

- 特别地，若$X_1, X_2, \ldots, X_n$为n个$i.i.d$随机变量，$\mathrm{Var}(X_i) = \sigma^2$，则

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

- $\mathrm{Var}(X_1 - X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$

7. $\mathrm{Var}(X) = 0$当且仅当$P(X = c) = 1$, 其中$c = E(X)$. 此时我们称X退化到常数c.
   对任何常数c，$\mathrm{Var}(X) \leq E(X - c)^2$, 等号成立当且仅当$c = E(X)$.

常见分布的方差

- 二项分布 $X \sim B(n,p)$
- Poisson分布 $X \sim P(\lambda)$
- 正态分布 $X \sim N(\mu,\sigma^2)$
- 均匀分布 $X \sim U(a,b)$
- 指数分布 $X \sim Exp(\lambda)$

- 我们称

$$Y = \frac{X - EX}{\sqrt{Var(X)}}$$

为随机变量X的标准化。

- $E(Y) = 0, \quad Var(Y) = 1.$

- 引入标准化随机变量是为了消除由于计量单位的不同而给随机变量带来的影响.
  - ✓ 例如, 考察人的身高, 可以以米为单位, 得到 X_1, 也可以以厘米为单位, 得到 X_2 .
  - ✓ 则 $X_2 = 100\, X_1$, X_2 与 X_1 的分布就有所不同. 不合理.
  - ✓ 通过标准化, 就可以消除两者之间的差别.

- **马尔可夫不等式：** 若随机变量$Y \geq 0$, 则对$\forall \varepsilon > 0$，有

$$P(Y \geq \varepsilon) \leq \frac{E(Y)}{\varepsilon}$$

- **切比雪夫不等式：** 若随机变量X的数学期望为$E(X) = \mu$, 则对$\forall \varepsilon > 0$，有

$$P(|X - \mu| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

例子

- 为了解某地区初中生人群沉迷于网络游戏的比例 p, 需要对该地区人群进行调查. 问要调查多少青少年, 才能使调查得到的网络游戏沉迷比例与真实网络游戏沉迷率之差的绝对值小于 0.5% 的概率不小于 95%.

3.2. 矩(Moment)

- 设X为随机变量，满足$E\left(|X|^k\right) < \infty$，$c$为常数，$k$为正整数，则

$$E\left[(X-c)^k\right]$$

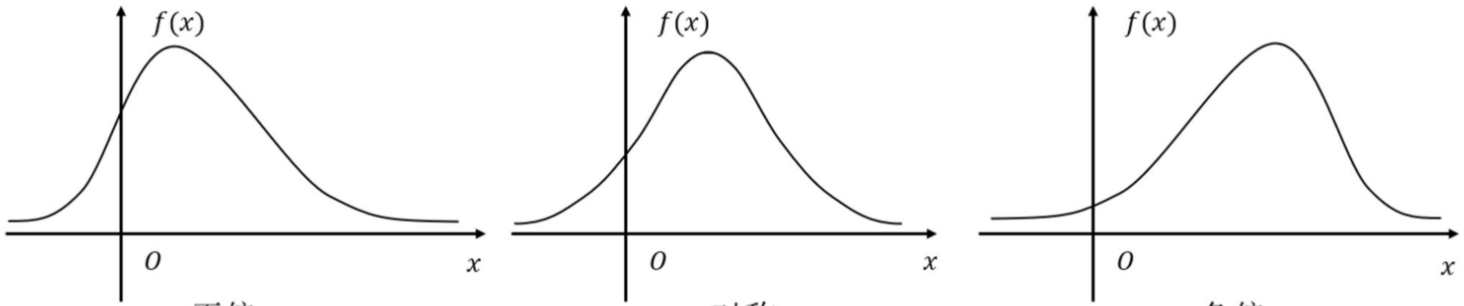
称为**X关于c点的k阶矩**。

- 两个重要的情况：
  - $c = 0$。这时 $\alpha_k = E\left(X^k\right)$ 称为**X的k阶原点矩**。
  - $c = E(X)$。这时 $\mu_k = E\left((X-EX)^k\right)$ 称为**X的k阶中心矩**。

- 期望$E(X) \triangleq \mu \Leftrightarrow X$的一阶原点矩$\alpha_1$；
- 方差$\mathrm{Var}(X) \triangleq \sigma^2 \Leftrightarrow X$的二阶中心矩$\mu_2$.
- **偏度系数(coefficient of skewness)：** 用三阶中心矩μ_3衡量分布是否有偏

正偏/右偏: $\mu_3 > 0$;　　对称: $\mu_3 = 0$;　　负偏/左偏: $\mu_3 < 0$

$$\boldsymbol{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{E(X-\mu)^3}{[E(X-\mu)^2]^{3/2}} = \boldsymbol{E\left(\frac{X-\mu}{\sigma}\right)^3} = \frac{\mu_3}{\sigma^3}$$（消除量纲影响）

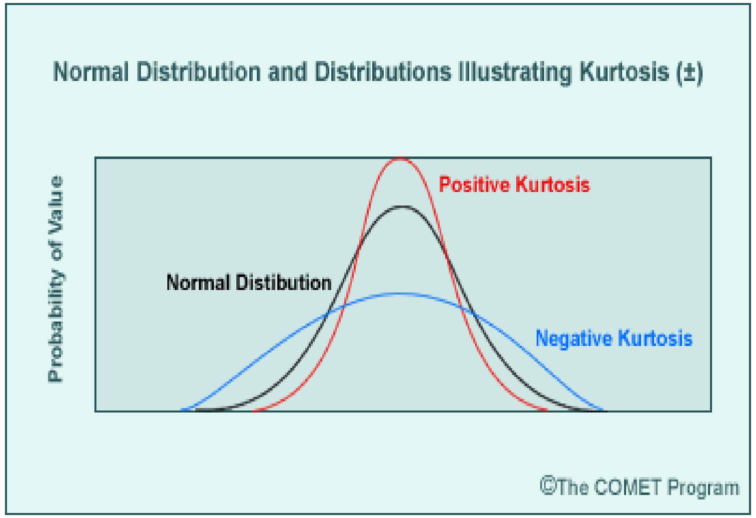


正偏 $(\mu_3 > 0)$　　对称 $(\mu_3 = 0)$　　负偏 $(\mu_3 < 0)$

- **峰度系数(coefficient of kurtosis)**：用四阶中心矩μ_4衡量分布(密度函数)在期望μ附近的陡峭程度

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{E(X-\mu)^4}{[E(X-\mu)^2]^2}=E\left(\frac{X-\mu}{\sigma}\right)^4=\frac{\mu_4}{\sigma^4}$$

或 $\beta_2=\dfrac{\mu_4}{\sigma^4}-\underset{\boxed{\text{正态}}}{\underset{\Uparrow}{3}}$



- **矩母函数(MGF):**

随机变量X的**矩母函数**或者矩生成函数 $M_X(s)$(moment generating function,简记为MGF) 定义为 $\boldsymbol{M_X(s)=E\left(e^{sX}\right)}$.

如果存在正常数a, 使得 $M_X(s)$对所有$s \in [-a, a]$是有限的，那么称X的矩母函数 $M_X(s)$存在.

$$M_X(s) = \mathbb{E}\left[e^{sX}\right] = \sum_{k=0}^{\infty} \mathbb{E}[X^k] \frac{s^k}{k!} \Rightarrow \boxed{\mathbb{E}X^k = \frac{d^k}{ds^k} M_X(s)\Big|_{s=0}}$$

- 设$X \sim N\left(\mu, \sigma^2\right)$，求$M_X(s)$.

- 矩母函数存在时可以唯一决定随机变量的分布

定理 4.2

假设存在正常数 c 使得随机变量 X 和 Y 的矩母函数对所有 $s \in [-c, c]$ 均有限且相等，则它们的分布相同. 即

$$F_X(t) = F_Y(t), \text{所有 } t \in \mathbb{R}.$$

- 独立随机变量的和的矩母函数等于各自矩母函数的乘积

随机变量 $X_1, X_2, \dots, X_n$相互独立，则

$$M_{X_1+X_2+\cdots X_n}(s) = M_{X_1}(s)\, M_{X_2}(s) \cdots M_{X_n}(s)$$

例子

- 设随机变量 X 的矩母函数为

$$M_X(s) = \frac{2}{2-s}, \qquad s \in (-2, 2)$$

证明其服从指数分布。