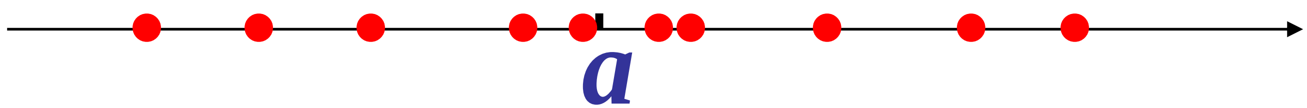
§ 4.2 方差Variance

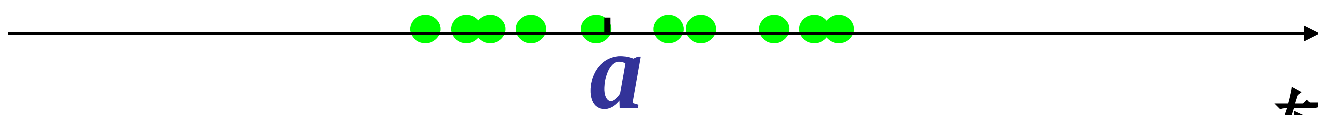
我们已经介绍了随机变量的数学期望，它体现了随机变量取值的平均水平，是随机变量的一个重要的数字特征.

但是在一些场合，仅仅知道平均值是不够的.

例如，某零件的真实长度为a，现用甲、乙两台仪器各测量10次，将测量结果X用坐标上的点表示如图：



甲仪器测量结果



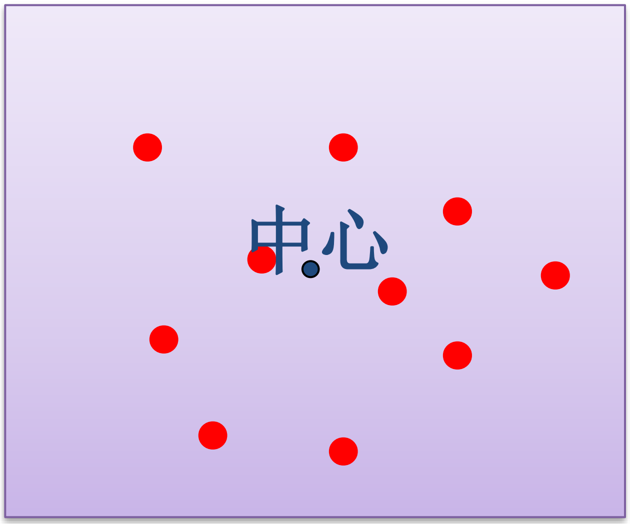
乙仪器测量结果

较好

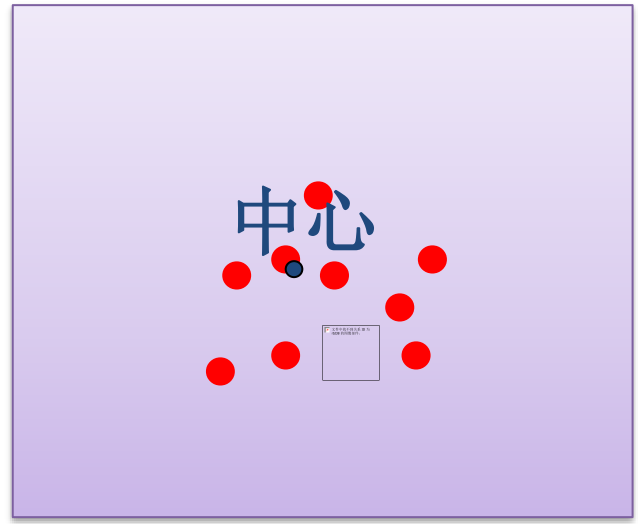
若让你就上述结果评价一下两台仪器的优劣，你认为哪台仪器好一些呢？

因为乙仪器的测量结果集中在均值附近

又如,甲、乙两门炮同时向一目标射击10发炮弹，其落点距目标的位置如图：



甲炮射击结果



乙炮射击结果

乙炮

你认为哪门炮射击效果好一些呢?

因为乙炮的弹着点较集中在中心附近.

为此需要引进另一个数字特征,用它来度量随机变量取值在其中心附近的离散Spread程度.

这个数字特征就是我们要介绍的

方差Variance

1. 方差概念

重要

定义 若$E\,[X - E(X)]^2$存在, 则称其为随机变量 X 的**方差**, 记为$D(X)$, $Var(X)$, $V(X)$, σ^2

即 $D(X)=E[X-E(X)]^2=\int[X-E(X)]^2dF(X)$

称 $\sqrt{D(X)}$ 为 X 的**均方差**或**标准差**Standard Variation.

$D(X)$ —— 描述 r.v. X 的取值偏离平均值的平均偏离程度

重要

由定义知，方差是随机变量X的函数 $g(X)=[X-E(X)]^2$的数学期望 .

若 X 为离散型 r.v.，分布律为

$$P(X = x_k) = p_k, \quad k = 1,2,\cdots$$

$$\boxed{D(X) = \sum_{k=1}^{+\infty} (x_k - E(X))^2 p_k}$$

若 X 为连续型r.v.，概率密度为 $f(x)$

$$\boxed{D(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx}$$

计算方差的一个简化公式

重要

$$D(X)=E(X^2)-E^2(X)$$

证：$D(X)=E[X-E(X)]^2$

展开

$$=E\{X^2-2XE(X)+[E(X)]^2\}$$

期望性质

$$=E(X^2)-2[E(X)]^2+[E(X)]^2$$

$$=E(X^2)-[E(X)]^2$$

例

X	-1	0	1
P	0.1	0.8	0.1

求$D(X)$.

$$E(X)=0,$$

$$E(X^2)=(-1)^2\cdot 0.1+1^2\cdot 0.1=0.2,$$

$$D(X)=E(X^2)-[E(X)]^2=0.2.$$

例 设随机变量X的密度函数为

$$f(x)=\begin{cases}1+x, & -1\le x\le 0\\ 1-x, & 0<x\le 1\end{cases}$$

求$D(X)$.

$$E(X)=\int_{-1}^{0}x(1+x)dx+\int_{0}^{1}x(1-x)dx=0,$$

$$E(X^2)=\int_{-1}^{0}x^2(1+x)dx+\int_{0}^{1}x^2(1-x)dx=\frac{1}{6},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{1}{6}$$

例 设$X \sim P(\lambda)$, 求$D(X)$.

解 $E(X)=\lambda$

$$\left.\begin{aligned} E(X^2) &= E(X(X-1))+E(X) \\ E(X(X-1)) &= \sum_{k=0}^{+\infty} k(k-1)\cdot\frac{\lambda^k e^{-\lambda}}{k!} \\ &= \lambda^2 e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!}=\lambda^2 \end{aligned}\right\} \Rightarrow E(X^2)=\lambda^2+\lambda$$

$$D(X)=E(X^2)-E^2(X)=\lambda$$

例 设 $X \sim U[a, b]$，求 DX.

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a \le x \le b, \\ 0, & \text{其他}\end{cases}$$

$$E(X)=\int_a^b x\cdot\frac{1}{b-a}dx=\frac{a+b}{2}$$

$$D(X)=E(X^2)-[E(X)]^2$$

$$=\int_a^b x^2\cdot\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}.$$

例 $X \sim E(\lambda)$，求 $D(X)$.

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, x>0\\ 0, \quad x\le 0\end{cases} \qquad E(X)=\frac{1}{\lambda}$$

$$D(X)=E(X^2)-[E(X)]^2$$

$$=\int_0^{+\infty} x^2\cdot\lambda e^{-\lambda x}dx-\left(\frac{1}{\lambda}\right)^2$$

$$=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

例 设 $X \sim N(\mu, \sigma^2)$, 求 $D(X)$

解 $E(X) = \mu$

$$D(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

注意方式

D(X)=E[X-E(X)]²=E([X-u]²)

$$\overset{\text{令}\frac{x-\mu}{\sigma}=t}{=} \int_{-\infty}^{+\infty} \sigma^2 t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= \sigma^2$$

常见随机变量的方差

重要

分布	概率分布	期望	方差
参数为p的0-1分布	$P(X=1)=p$ $P(X=0)=1-p$	p	$p(1-p)$
$B(n,p)$	$P(X=k)=C_n^k p^k (1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np	$np(1-p)$
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\cdots$	λ	λ

重要

分布	概率密度	期望	方差
区间(a,b)上的均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & 其它\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & 其它\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

2. 方差的性质

重要

(1) 设C是常数,则$D(C)=0$;

(2) 若C是常数,则$D(CX)=C^2 D(X)$;

(3) 若X与Y独立，则

$$D(X+Y)=D(X)+D(Y).$$

X与Y不一定独立时，

$D(X+Y)=?$

D(X-Y)?

推广：若$X_1,X_2,\ldots,X_n$相互独立,则

$$D\left[\sum_{i=1}^{n} X_i\right]=\sum_{i=1}^{n} D(X_i)$$

$$D\left[\sum_{i=1}^{n} C_i X_i\right]=\sum_{i=1}^{n} C_i^2 D(X_i)$$

例 设$X \sim B(n, p)$，求$D(X)$.

解一 利用定义$D(X)=E([X-E(X)]^2)$

解二 利用简化公式求$D(X)=E(X^2)-E^2(X)$.

解三 利用性质求$D(X)$.

引入随机变量 $X_1, X_2, \cdots, X_n$

$$X_i = \begin{cases} 1, & \text{第}\, i \,\text{次试验事件}\, A \,\text{发生} \\ 0, & \text{第}\, i \,\text{次试验事件}\, \bar{A} \,\text{发生} \end{cases}$$

$$D(X_i) = p(1-p) \qquad i = 1, 2, \cdots, n$$

$X_1, X_2, \cdots, X_n$ 相互独立， $\quad X = \sum_{i=1}^{n} X_i$

故 $$D(X) = \sum_{i=1}^{n} D(X_i) = np(1-p)$$

Summary

1方差定义

2方差的简化公式

3常见分布的方差

4方差性质

5 利用1、2、4进行计算