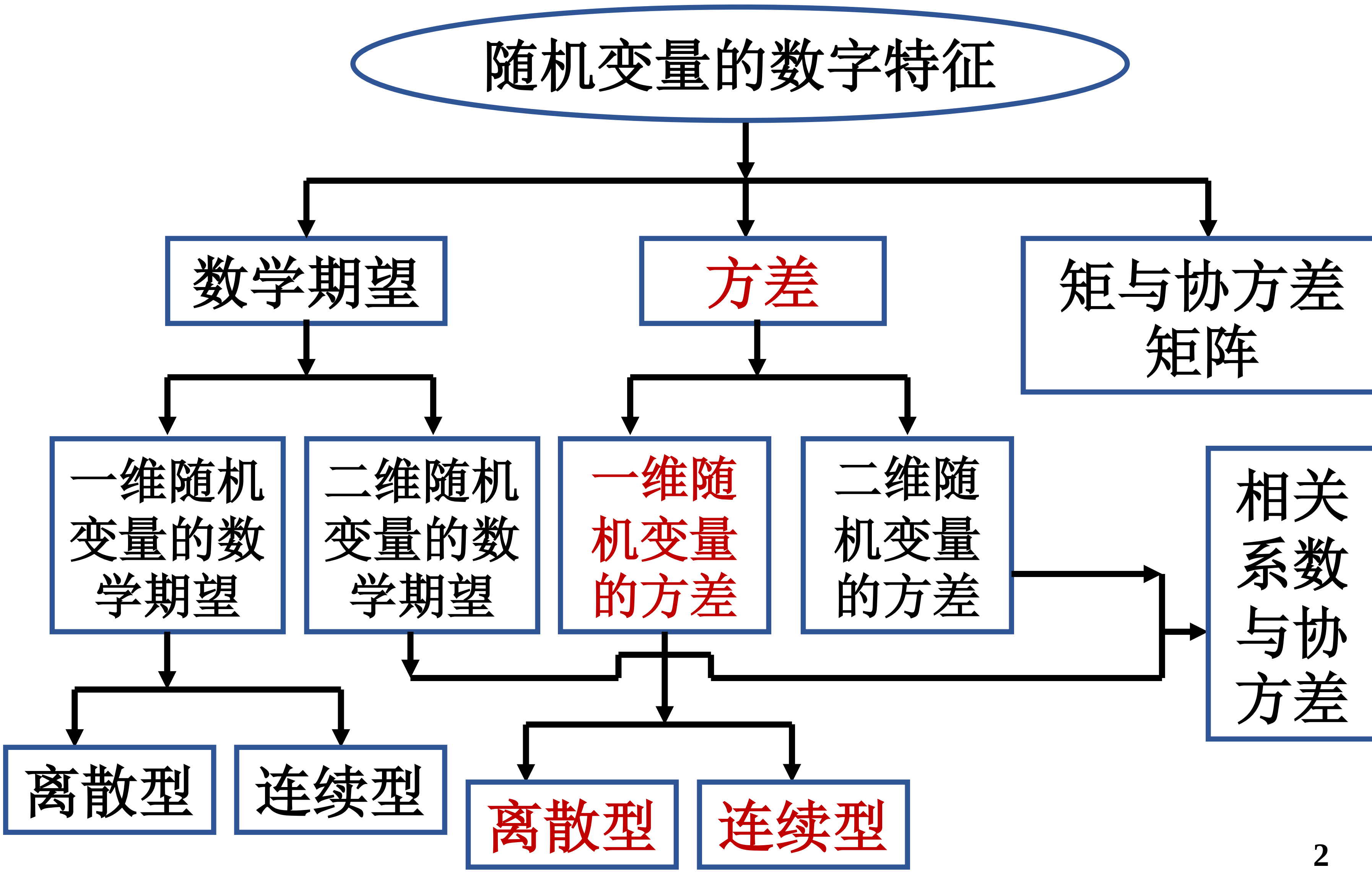
第四章
随机变量的数字特征
—刻画随机变量某一方面特征的常数

王笑尘

北京邮电大学网络空间安全学院

wxiaochen@bupt.edu.cn

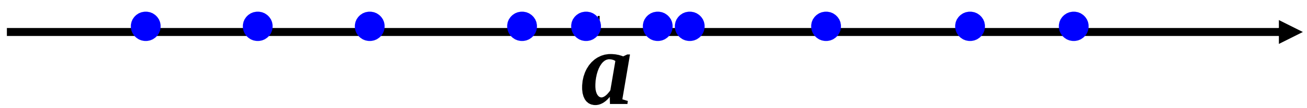
第二节 方差


随机变量的数字特征
数学期望
方差
矩与协方差矩阵
一维随机变量的数学期望
二维随机变量的数学期望
一维随机变量的方差
二维随机变量的方差
相关系数与协方差
离散型
连续型
离散型
连续型

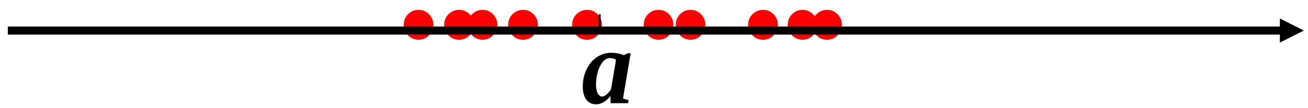
第二节 方差

问题的引出

引例1 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：

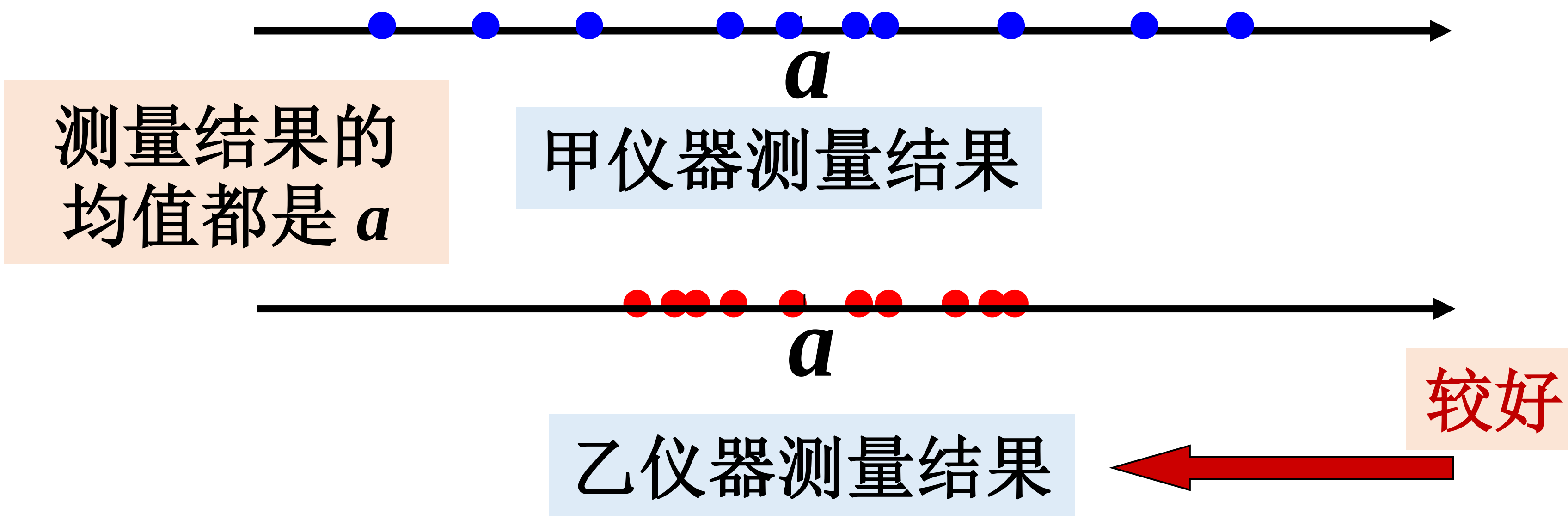


甲仪器测量结果

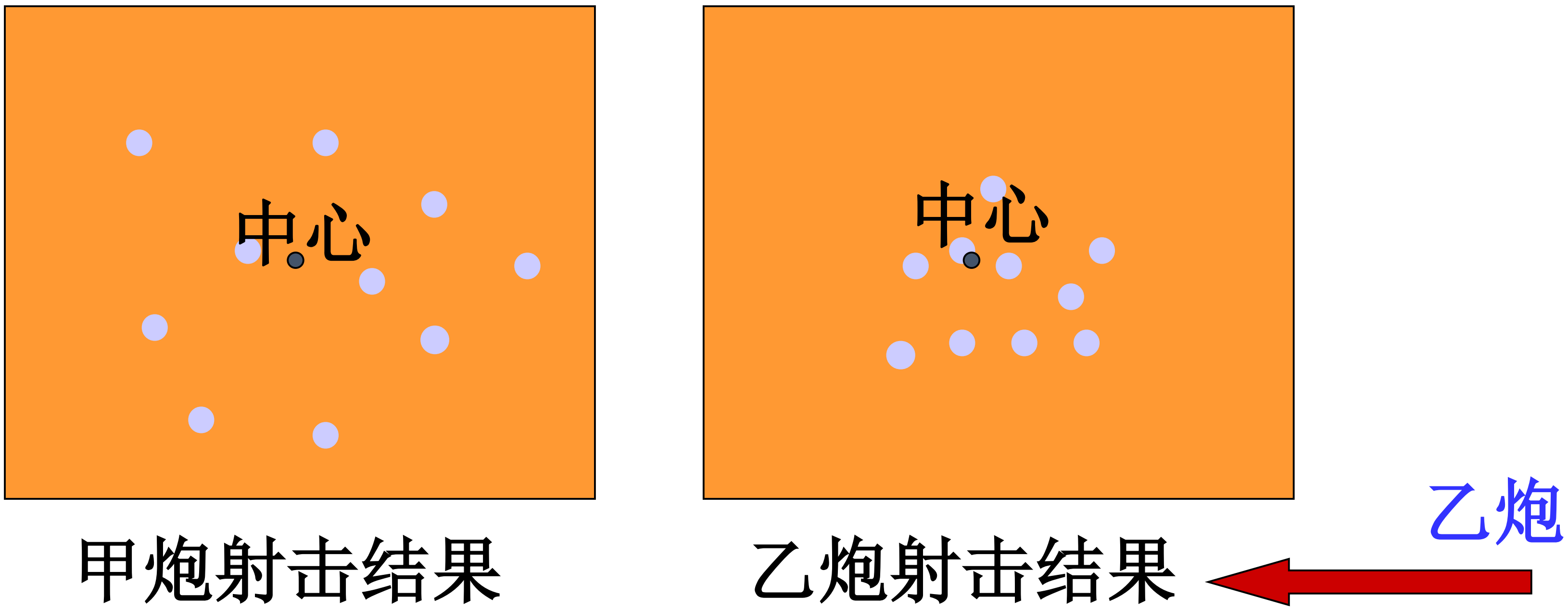


乙仪器测量结果

因为乙仪器的测量结果集中在均值附近，若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？



引例2 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：



哪门炮射击效果好一些呢?

因为乙炮的炮弹着点较集中在中心附近。

引例3 某两个储蓄所，它们的月吸收存款额(万元)及其概率如下所示：

甲储蓄所

月吸收存款额 $X_{甲}$	8	10	12
概率 p	0. 2	0. 6	0. 2

乙储蓄所

月吸收存款额 $X_{乙}$	7	10	13
概率 p	0. 3	0. 4	0. 3

问：甲乙两个储蓄所哪个月吸收存款额更稳定?

解：(1) 若计算其数学期望，则：

$$E(X_{甲}) = 8 \times 0.2 + 10 \times 0.6 + 12 \times 0.2 = 10\,(万元)$$

$$E(X_{乙}) = 7 \times 0.3 + 10 \times 0.4 + 13 \times 0.3 = 10\,(万元)$$

- 从计算的结果上看，这两个储蓄所的月吸收存款额的平均值是相同的。
- 但从已知条件稍加分析可知，甲储蓄所的月吸收存款额比乙储蓄所的月吸收存款额来得“稳定”。因此，需进一步量化它的取值与平均值的偏离程度。

(2) 若用随机变量与其数学期望的偏差的期望值来表示这偏离程度，则：

$$E[X_{甲}-E(X_{甲})]=\boxed{(8-10)\times 0.2}+(10-10)\times 0.6$$
$$+(12-10)\times 0.2=0\ (万元)$$

$$E[X_{乙}-E(X_{乙})]=\boxed{(7-10)\times 0.3}+(10-10)\times 0.4$$
$$+(13-10)\times 0.3=0\ (万元)$$

➢ 从计算的结果上看，由于诸偏差的正负抵消，这两个储蓄所的月吸收存款额与其数学期望的偏差的期望值均为“0”，这样就掩盖了实际偏差的大小。

(3) 因此，为了克服诸偏差的正负抵消，真正反映出实际偏差的大小程度，通常采用随机变量与其数学期望的偏差平方的期望值来表示这偏离程度，则：

$$E\{[X_{甲}-E(X_{甲})]^2\}=(8-10)^2\times 0.2+(10-10)^2\times 0.6$$
$$+(12-10)^2\times 0.2=1.6 \quad (万元)$$
$$E\{[X_{乙}-E(X_{乙})]^2\}=(7-10)^2\times 0.3+(10-10)^2\times 0.4$$
$$+(13-10)^2\times 0.3=5.4 \quad (万元)$$

- 结果显示：偏差平方的期望值真正反映出实际偏差的大小程度：甲储蓄所的月吸收存款额比乙储蓄所的“稳定”。
- 通常称用偏差平方的数学期望来描述随机变量的取值与平均值的偏离程度为“方差”。

一. 方差的定义

定义. 设X是一个随机变量，若 $E\{[X-E(X)]^2\}$ 存在，则称 $D(X)=E\{[X-E(X)]^2\}$ 为 X 的方差。记为：

$$D(X)=Var(X)=E\{[X-E(X)]^2\}$$

采用平方是为了保证一切差值$X-E(X)$都起正面的作用

注 ▲ $D(X)$是个(实)数，它形式上是X 的每一个取值和它们的平均值的偏差平方与相应概率的乘积之和；

▲ 本质上体现了X 围绕着“平均值”的偏离程度，故它是衡量X 取值分散程度的一个标志。

▲ 方差的算术平方根 $\sqrt{D(X)}$ 称为标准差或均方差。记为:

$$\sigma(X)=\sqrt{D(X)}$$

▲ $D(X)$实际上是X的函数 $g(X)=(X-E(X))^2$ 的数学期望。

二. 离散型随机变量的方差

定义. 设离散型随机变量X 的分布律为:

X	x_0	x_1	x_2	$\cdots\cdots$	x_n	$\cdots\cdots$
p_k	p_0	p_1	p_2	$\cdots\cdots$	p_n	$\cdots\cdots$

如果级数 $\sum_{k=1}^{\infty}[x_k - E(X)]^2 p_k$ 绝对收敛，则称此级数为 X 的方差，记为:

$$D(X) = Var(X) = \sum_{k=1}^{\infty}[x_k - E(X)]^2 p_k = E\{[X - E(X)]^2\}$$

注 如果级数 $\sum_{k=1}^{\infty}[x_k - E(X)]^2 p_k$ 不绝对收敛，则称 $D(X)$ 不存在。

三．连续型随机变量的方差

定义． 设连续型随机变量X 的概率密度为 $f(x)$ 如果积分 $\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx$ 绝对收敛，则称此积分为 X 的方差，记为：

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

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

四. 方差的性质

这是一个重要的经常使用的计算公式

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

证明: $\because \quad D(X)=E\{[X-E(X)]^2\}$

由数学期望的性质

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

因为数学期望E(X)是数

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

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

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

注: 这个公式给出了计算随机变量X的方差的公式，同时也给出了数学期望与方差之间的关系。

2．若 c 是常数，则：$D(c)=0$

3. 若 C 是常数，X 是随机变量，则：

$$D(cX)=c^2D(X)$$

证明：$D(cX)=E[(cX)^2]-[E(cX)]^2$

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

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

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

$D(X+C)=D(X)$

X+C相对于X，即取值整体平移，所以只改变均值位置（大小），而不改变数据整体的分散程度。

4. 若X,Y是相互独立的随机变量，则：

$D(X+Y)=D(X)+D(Y)$　证明见后

证明： $D(X+Y)=E\{[(X+Y)-E(X+Y)]^2\}$ 由方差定义

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

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

$$+2\boxed{E\{[X-E(X)][(Y-E(Y)]\}}$$

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

$$+2E[X-E(X)]\cdot E[Y-E(Y)]$$

因为X,Y相互独立，所以X-E(X)与Y-E(Y)也相互独立

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

$$+2[E(X)-E(X)]\cdot[E(Y)-E(Y)]$$

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

注：此性质可推广到任意有限多个相互独立的随机变量之和的情形。

定义和性质应用—离散型随机变量几种常见分布的方差

(1) $(0-1)$ 分布

若随机变量 X 只能取 0 与 1 两个值，它的分布律为:

$$P(X=k)=p^k(1-p)^{(1-k)} \qquad k=0,1. \quad 0<p<1$$

则： $\mathrm{E(X)}=p$

$$\mathrm{D(X)}=(0-p)^2\cdot q+(1-p)^2\cdot p=pq$$

(2) 二项分布

设随机变量 X 服从参数为(n,p)的二项分布，

即 $X \sim B(n, p)$，它的分布律为:

$$P(X=k)=C_n^k p^k (1-p)^{n-k}, \quad k=0,1,2\cdots n$$

则： $\boxed{E(X)=np}$

$$\boxed{D(X)=npq}$$

证明见后

例1 设 $X_1, X_2, \cdots, X_n$ 相互独立，且服从同一 $(0-1)$ 分布

1. 证明：$X = X_1 + X_2 + \cdots + X_n$ 服从参数为 n, p 的二项分布

2. 求：$E(X),\ D(X)$

解: 1. 因为 X 的所有取值为：$0, 1, 2, \cdots, n$

又因为：X_i 相互独立 $(i = 1, 2, \cdots, n)$

所以X 取 k $(0 \le k \le n)$的概率为:

故有: $P(X = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, 2, \cdots, n$

即X 服从参数为 n, p 的二项分布。

参数为(n, p)的二项分布变量，可分解为n个相互独立且都服从以p为参数的（0-1）分布的随机变量之和。

2. 因为 X_i 服从(0-1)分布 $(i = 1, 2, \cdots, n)$

所以有:

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

而二项分布变量X:

$$E(X) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = np$$

又由已知，$X_1, X_2, \cdots, X_n$ 相互独立

所以有:

$$D(X) = D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D(X_i) = npq$$

(3) 泊松分布

若随机变量X 的所有可能取值为: $0, 1, 2, \cdots\cdots$
而它的分布律(它所取值的各个概率)为:

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \quad k=0,1,2,\cdots \qquad \text{即:} X \sim \pi(\lambda)$$

则: $\mathrm{E(X)}=\lambda$ $\quad$ $\mathrm{D(X)}=\lambda$

定义和性质应用—连续型随机变量几种常见分布的方差

(1). 均匀分布

若连续型随机变量 X 具有概率密度$f(x)$ 为:

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

即 $X\sim U[a,b]$

则：

$$D(X)=\frac{(b-a)^2}{12}$$

$$E(X)=\frac{a+b}{2}$$

证明见后

例2. 设 $X \sim U[a,b]$, 即它的概率密度函数为：

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

求：X 的方差

解：因为：$D(X)=E(X^2)-[E(X)]^2$

$$E(X)=\frac{a+b}{2}$$

所以：$D(X)=\int_a^b x^2\frac{1}{b-a}dx-(\frac{a+b}{2})^2$

$$=\frac{(b-a)^2}{12}$$

(2). 指数分布

若连续型随机变量 X 具有概率密度$f(x)$为:

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & 其它\end{cases}$$

其中$\theta>0$为常数

则: $$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx$$

$$=\int_{0}^{+\infty}(x-\theta)^2\cdot\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=\theta^2$$

即: $$D(X)=\theta^2$$

$$E(X)=\theta$$

(3). 正态分布

若随机变量 X 的概率密度为：

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty$$

即：$X\sim N(\mu,\sigma^2)$

则：$D(X)=\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)dx$

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

令：$z=\dfrac{x-\mu}{\sigma}\quad\Rightarrow$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma^2 z^2\frac{1}{\sigma}e^{-\frac{z^2}{2}}\sigma\,dz=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}z^2 e^{-\frac{z^2}{2}}dz$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left[-ze^{-\frac{z^2}{2}}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz\right]=\frac{\sigma^2}{\sqrt{2\pi}}\cdot\sqrt{2\pi}=\sigma^2$$

即: $\boxed{D(X)=\sigma^2}$ $\boxed{E(X)=\mu}$

结论：正态分布中密度函数的参数 σ^2 恰好就是随机变量 X 的方差。

正态分布中密度函数的参数 μ 恰好就是随机变量X的数学期望。

例3. 设随机变量 X 服从几何分布，其分布律为：

$$P(x=k)=p(1-p)^{k-1},k=1,2,\cdots$$

其中，$0<p<1$

求：$D(X)$

例3. 设随机变量 X 服从几何分布，其分布律为：

$$P(x=k)=p(1-p)^{k-1}, k=1,2,\cdots$$

其中，$0<p<1$

求：$D(X)$

解：记 $q=(1-p)$

$$E(X)=\boxed{\sum_{k=1}^{\infty}kpq^{k-1}=p\sum_{k=1}^{\infty}(q^k)'}$$

$$=p\left(\sum_{k=1}^{\infty}q^k\right)'=p\left(\frac{q}{1-q}\right)'=\frac{1}{p}$$

求和与求导交换次序

无穷递缩等比级数求和公式

$$E(X^2)=\sum_{k=1}^{\infty}k^2pq^{k-1}$$ $P(x=k)=p(1-p)^{k-1}, k=1,2,\cdots$

$$=p[\sum_{k=1}^{\infty}k(k-1)q^{k-1}+\sum_{k=1}^{\infty}kq^{k-1}]$$

$$=qp(\sum_{k=1}^{\infty}q^k)''+E(X)=qp(\frac{q}{1-q})''+\frac{1}{p}$$

$$=qp\frac{2}{(1-q)^3}+\frac{1}{p}=\frac{2q}{p^2}+\frac{1}{p}=\frac{2-p}{p^2}$$

所以: $D(X)=E(X^2)-[E(X)]^2$

$$=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}$$

5. $D(X)=0$ 的充分必要条件是X以概率1取常数 c

即 $P(X=c)=1$ （后面证）

6. (切比雪夫不等式) 设随机变量X具有数学期望 $E(X)=\mu$, 方差 $D(X)=\sigma^2$. 则对任意正数 ε

不等式: $P\{|X-\mu|\geq\varepsilon\}\leq\dfrac{\sigma^2}{\varepsilon^2}$ 成立。

称其为切比雪夫不等式。

切比雪夫不等式(chebysev)的另一形式:

$$P(|X-\mu|<\varepsilon)\geq 1-\frac{\sigma^2}{\varepsilon^2}$$

证明：就连续型随机变量的情况来证明

设X的概率密度为$f(x)$,则有

$$P(|X-\mu|\geq\varepsilon)=\int_{|x-\mu|\geq\varepsilon}f(x)dx$$

$$\leq\int_{|x-\mu|\geq\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2f(x)dx$$

$$=\frac{\sigma^2}{\varepsilon^2}$$

- 切比雪夫不等式的作用：给出了在随机变量X的分布未知的情况下，$\{|X-\mu|\geq\varepsilon\}$ 概率的上限的一种估计方法。

- 如取 $\varepsilon=3\sigma$，则:

$$P\{|X-E(X)|\geq 3\sigma\}\leq\frac{\sigma^2}{9\sigma^2}\approx 0.111$$

可见，对任给的分布，只要期望和方差 σ^2 存在，则随机变量 X 取值偏离 $E(X)$ 超过 3σ 的概率小于0.111 。

$$P\{|X-\mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

由切比雪夫不等式可以看出，若 σ^2越小，则事件 $\{|X-E(X)|<\varepsilon\}$的概率越大，即随机变量 X 集中在期望附近的可能性越大。由此可体会方差的概率意义

它刻划了随机变量取值的离散程度

现应用切比雪夫不等式证明性质5：

$D(X)=0$的充分必要条件是 X 以概率 1 取常数c
即 $P(X=c)=1$

证明：$\Leftarrow \because\ P(X=c)=1 \quad \therefore\ D(X=c)=0$

$\Rightarrow \quad \because\ \boxed{D(X)=0}$

所以由切比雪夫不等式可知：

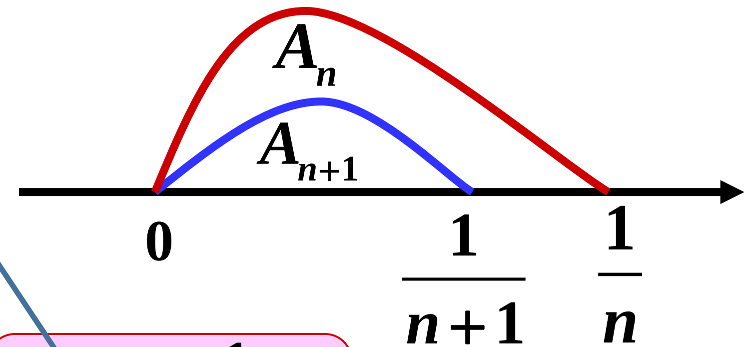
$$0 \le P\{|X-\mu| \ge \varepsilon\} \le \frac{D(X)}{\varepsilon^2} = 0$$

即：$P\{|X-\mu| \ge \varepsilon\} = 0$

A_n 事件是在0与 $1/n$之间

则其对立事件：$P\{|X-\mu|<\varepsilon\}=1$

取 $\varepsilon=\dfrac{1}{n}$，记 $A_n=\{|X-E(X)|<\dfrac{1}{n}\}$



显然，$A_n\supset A_{n+1}$

令：

$$A=\bigcap_{n=1}^{\infty}A_n=\{|X-E(X)|=0\}$$
$$=\{X=E(X)\}$$

$\because\ \varepsilon=\dfrac{1}{n}$

由连续性得：

$$P(A)=\lim_{n\to\infty}P(A_n)=\lim_{n\to\infty}P(|X-E(X)|<\frac{1}{n})=1$$

即：$P(X=E(X))=1$　显然，$c=E(X)$

另：反证法（见教材P107）

例4 已知正常男性成人血液中，每一毫升白细胞数平均是7300，均方差是700 。利用切比雪夫不等式估计每毫升白细胞数在5200 ~ 9400之间的概率 。

例4 已知正常男性成人血液中，每一毫升白细胞数平均是7300，均方差是700 。利用切比雪夫不等式估计每毫升白细胞数在5200 ~ 9400之间的概率 。

解： 设每毫升白细胞数为 X

依题意，$E(X) = 7300, \quad D(X) = 700^2$

现求： $P(5200 \le X \le 9400) = ?$

$$P(5200 - 7300 \le X - 7300 \le 9400 - 7300)$$

$$= P(-2100 \le X - E(X) \le 2100)$$

$$= P(|X - E(X)| \le 2100)$$

由切比雪夫不等式得：

$$P\{|X-E(X)|\le 2100\}\ge 1-\frac{D(X)}{(2100)^2}$$

$$=1-(\frac{700}{2100})^2 \quad =1-\frac{1}{9}=\frac{8}{9}$$

即估计每毫升白细胞数在5200 ~ 9400之间的概率不小于8/9 。

例5 在每次试验中，事件A 发生的概率为 0.75, 利用切比雪夫不等式求：n 需要多么大时，才能使得在 n 次独立重复试验中, 事件 A 出现的频率在0.74 ~ 0.76之间的概率至少为 0. 90 ?

自行考虑

解： 设 X 为 n 次试验中，事件 A 出现的次数

则：$X \sim B(n,\ 0.75)$

$$E(X)=0.75n \qquad D(X)=0.75\times 0.25n=0.1875n$$

所求为满足： $P(0.74\le \frac{X}{n}\le 0.76)\ge 0.9$ 的最小的 n

$P(0.74\le \frac{X}{n}\le 0.76)$ 可改写为：

$$P(0.74n\le X\le 0.76n)$$

$$=P(0.74n-0.75n\le X-0.75n\le 0.76n-0.75n)$$

$$=P(-0.01n\le X-0.75n\le 0.01n)$$

$$=P(\left|X-E(X)\right|\le 0.01n)$$

在切比雪夫不等式中取 $\varepsilon = 0.01n$：

则有：

$$P(0.74 \le \frac{X}{n} \le 0.76) = P(\left|X - E(X)\right| \le 0.01n)$$

$$\ge 1 - \frac{D(X)}{(0.01n)^2} = 1 - \frac{0.1875n}{0.0001n^2} = 1 - \frac{1875}{n}$$

依题意，取：$1 - \frac{1875}{n} \ge 0.9$

解得：$n \ge \frac{1875}{1 - 0.9} = 18750$

即 n 取18750 时，可以使得在 n 次独立重复试验中, 事件 A 出现的频率在0.74 ~ 0.76之间的概率至少为0.90。

例6 设 X 的数学期望为 E(X)，方差为 D(X)

求：随机变量 $Y=\dfrac{X-E(X)}{\sqrt{D(X)}}$ 的数学期望与方差

例6 设 X 的数学期望为 E(X)，方差为 D(X)

求：随机变量 $Y=\dfrac{X-E(X)}{\sqrt{D(X)}}$ 的数学期望与方差

解: $E(Y)=E(\dfrac{X-E(X)}{\sqrt{D(X)}})=\dfrac{1}{\sqrt{D(X)}}E[X-E(X)]$

$$=\frac{1}{\sqrt{D(X)}}[E(X)-E(X)]=0$$

$$D(Y)=D(\frac{X-E(X)}{\sqrt{D(X)}})=\frac{1}{D(X)}D[X-E(X)]$$

$$=\frac{D(X)}{D(X)}=1$$

$$D(X+C)=D(X)$$

称 Y 为标准化随机变量