复习

数学期望
E(X)

离散型

连续型

$$E(X)=\sum_{k=1}^{\infty}x_k p_k$$

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

随机变量函数的数学期望

$$\sum_{k=1}^{\infty}g(x_k)p_k$$

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

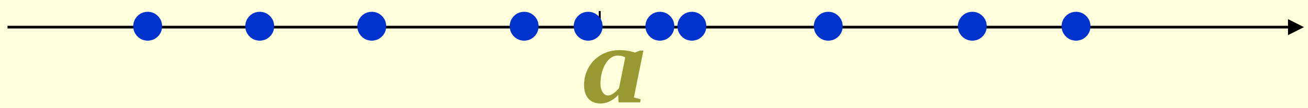
期望性质

§2 方 差

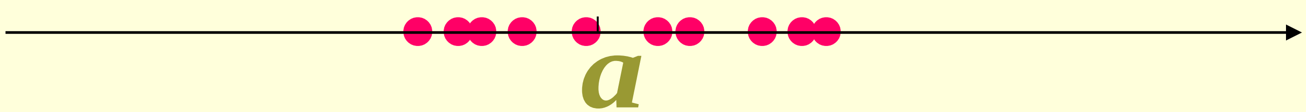
上一讲我们介绍了随机变量的数学期望，它体现了随机变量取值的平均水平，是随机变量的一个重要的数字特征.

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：



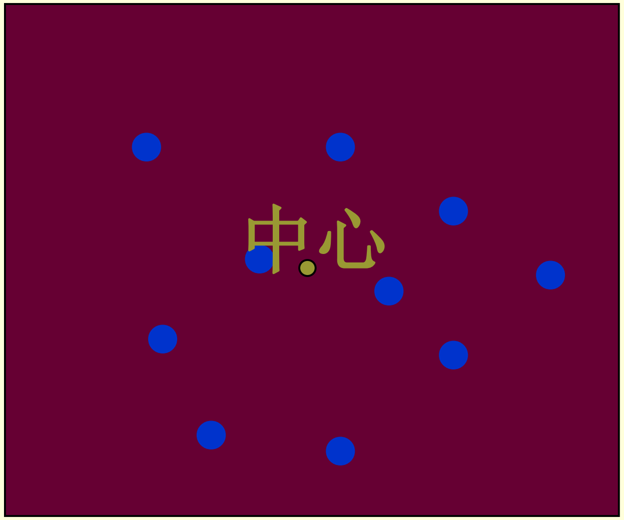
甲仪器测量结果



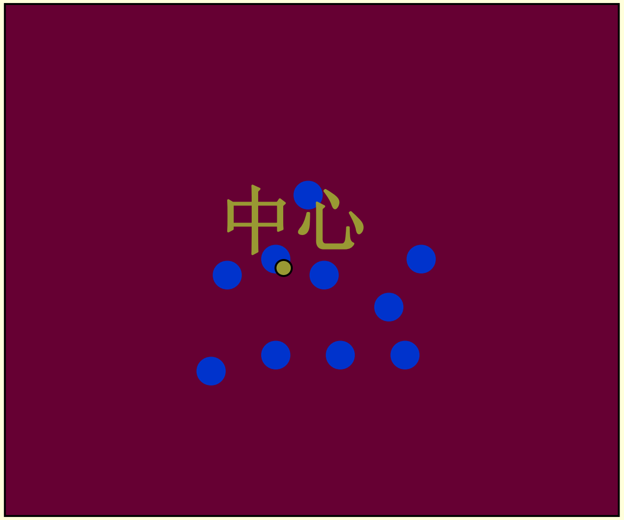
乙仪器测量结果

测量结果的均值都是 a

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

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相对于其均值的离散程度.

这个数字特征就是我们这一讲要介绍的

方差

方差

设X是一个随机变量，若$E[(X-E(X)]^2<\infty$，则称

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

为X的方差.

方差的算术平方根$\sqrt{Var(X)}$称为**标准差**

注

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

(1) $Var(X)\geq 0$，即方差是一个非负实数

(2)当X服从某分布时，我们也称某分布的方差为$Var(X)$。

(3)方差刻划了随机变量的取值对于其数学期望的离散程度。

若X的取值比较集中，则方差较小；

若X的取值比较分散，则方差较大．

方差的计算公式

(1) 若X为离散型随机变量，其分布律为

$$p_i = P(X=x_i),\ i=1, 2, \ldots ,$$

且$Var(X)$存在，则

$$\mathrm{Var}(X) = \sum_i (x_i - E(X))^2 p_i$$

若X为连续型随机变量，其概率密度为

$f(x)$，且$D(X)$存在，则

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx$$

（2）若随机变量的方差*Var* (*X*)存在，

则

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

证： $Var(X)=E[X-E(X)]^2$

展开

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

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

利用期望性质

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

三、常见分布的方差

例1 设$X \sim B(1, p)$，求$Var(X)$

例2 设$X \sim B(n,p)$，求$Var(X)$

例3 设$X \sim \pi(\lambda)$，求$Var(X)$

例4 设$X \sim U[a,b]$，求$Var(X)$

例5 设X服从参数为λ的指数分布，求$Var(X)$。

例6 设$X \sim N(\mu, \sigma^2)$，求$Var(X)$

例1　设$X\sim B(1, p)$，求$Var(X)$

$$E(X)=P(X=1)=p, \qquad E(X^2)=p,$$

故　$Var(X)=E(X^2)-[E(X)]^2=p-p^2=p(1-p)$

例2 设$X\sim B(n,p)$，求$Var(X)$

解：$EX^2 = E(X(X-1)+X)$

只需求E(X(X-1))

$$E(X(X-1)) = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!} p^k (1-p)^{n-k}$$

$$= n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{n-2-(k-2)}$$

二项式定理

$$= n(n-1)p^2$$

所以 $EX^2 = E(X(X-1)+X) = n(n-1)p^2 + np$

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = np(1-p)$$

例3 设$X \sim \pi(\lambda)$, 求$Var(X)$.

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2$$

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

解 $$E(X) = \lambda$$

$$E(X^2) = E(X(X-1)) + E(X)$$

$$E(X(X-1)) = \sum_{k=0}^{+\infty} k(k-1) \cdot \frac{\lambda^k e^{-\lambda}}{k!}$$

$$= \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

$$\Rightarrow E(X^2) = \lambda^2 + \lambda$$

$$\mathrm{Var}(X) = E(X^2) - E^2(X) = \lambda$$

例5 **设X服从参数为λ的指数分布，求$Var(X)$。**

解：EX^2

$$\int_0^{\infty} x\cdot\lambda e^{-\lambda x}dx = 1/\lambda$$

$$= \int_0^{\infty} \lambda x^2 e^{-\lambda x}dx$$

$$= -\int_0^{\infty} x^2 de^{-\lambda x}$$

$$= -[x^2 e^{-\lambda x}\Big|_0^{\infty} - 2\int_0^{\infty} x e^{-\lambda x}dx] \quad = \frac{2}{\lambda^2}$$

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = 1/\lambda^2$$

例6 设 $X \sim N(\mu, \sigma^2)$, 求 $Var(X)$

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

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

$$= \sigma^2$$

常见随机变量的方差

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

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

例7 设r.v X服从几何分布，概率函数为

$$P(X=k)=p(1-p)^{k-1},\ k=1,2,\ldots$$

其中$0<p<1$,求$Var(X)$

解： 记$q=1-p$

无穷递缩等比级数求和公式

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

求和与求导交换次序

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

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

$$E(X(X-1))$$

$$= p\sum_{k=1}^{\infty} k(k-1)q^{k-1} = pq\sum_{k=1}^{\infty} k(k-1)q^{k-2}$$

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

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

$$E(X^2) = \frac{2q}{p^2}+\frac{1}{p} = \frac{2-p}{p^2}$$

$$\mathbf{Var\,(X)=E(X^2)-[E(X)]^2} = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

例8 设X 的概率密度为

$$f(x) = ae^{-\frac{(x-3)^2}{8}}$$

a为未知常数, 求a, $E(X^2)$.

例9

设$X,Y \sim U[0,1]$, 且相互独立。**求**$\boldsymbol{Var}\,(\min\{X,Y\}$

法一

解:

$$F_X(x) = F_Y(x)$$

$$= \begin{cases} 0 & x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$

设$Z=\min(X,Y)$，Z的分布函数为

$$F_X(x)=F_Y(x)=\begin{cases}0 & x\le 0\\ x & 0<x<0\\ 1 & x\ge 1\end{cases}$$

$$F_Z(z)=1-[1-F_X(z)][1-F_Y(z)]$$

$$=\begin{cases}0 & z\le 0\\ 1-(1-z)^2 & 0<z<1\\ 1 & z\ge 1\end{cases}$$

其密度函数为

$$f_Z(z)=F_Z{}'(z)=\begin{cases}2(1-z) & 0<z<1\\ 0 & 其他\end{cases}$$

$$E(Z)=\int_0^1 2z(1-z)dz==\frac{1}{3}$$

$$f_Z(z)=\begin{cases}2(1-z) & 0<z<1\\ 0 & \text{其他}\end{cases}$$

$$E(Z^2)=\int_0^1 2z^2(1-z)dz=\frac{1}{6}$$

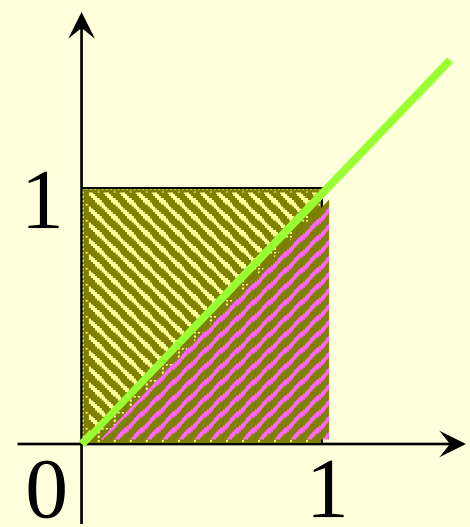
$$\mathrm{Var}(Z)=E(Z^2)-[E(Z)]^2=\frac{1}{18}$$

在 [0, 1] 中随机地取两个数 *X* , *Y* , 求 *Var* (min{ *X* ,*Y* })

法二 $\mathrm{Var}(\min\{X,Y\})$

$$= E\left(\min{}^2\{X,Y\}\right) - E^2\left(\min\{X,Y\}\right)$$

解 $$f(x,y) = \begin{cases} 1, & 0<x<1, 0<y<1 \\ 0, & \text{其它} \end{cases}$$



$$E(\min\{X,Y\})$$

$$= \iint\limits_{\substack{0<x<1\\0<y<1}} \min\{x,y\}\,dxdy$$

$$= \int_0^1\left(\int_x^1 x\,dy\right)dx + \int_0^1\left(\int_y^1 y\,dx\right)dy \quad = \frac{1}{3}$$

$$E(\min{}^2\{X,Y\})$$

$$=\int_0^1\left(\int_x^1 x^2dy\right)dx+\int_0^1\left(\int_y^1 y^2dx\right)dy$$

$$=\frac{1}{6}$$

$$\mathrm{Var}(\min\{X,Y\})$$

$$=E\left(\min{}^2\{X,Y\}\right)-E^2\left(\min\{X,Y\}\right)$$

$$=\frac{1}{18}$$

方差的性质

性质1 若$X=c$，c为常数，则 $Var(X)=0$

性质2 若c为常数,随机变量X的方差存在，则cX的方差存在，且

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

推论： $Var(cX+b) = c^2 Var(X)$

性质 1 的证明：

$$\mathrm{Var}(C)=E\big(C-E(C)\big)^2=0$$

性质 2 的证明：

$$\begin{aligned}\mathrm{Var}(cX)&=E\big(cX-E(cX)\big)^2\\&=E\big(cX-cE(X)\big)^2\\&=E\big(c^2(X-E(X))^2\big)\\&=c^2\mathrm{Var}(X)\end{aligned}$$

性质3 若$D(X)$, DY 存在，则

$$Var(X\pm Y)= Var(X)+ Var(Y) \pm 2E[(X-E(X))(Y-E(Y))]$$

若随机变量X, Y相互独立，它们的方差都存在，则$X\pm Y$的方差也存在，且

$$Var(X\pm Y)=Var(X)+Var(Y)$$

推论 若随机变量$X_1,X_2,\dots,X_n$相互独立，它们的方差都存在，则$a_1X_1+a_2X_2+\dots+a_nX_n$的方差存在，且

$$\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right)=\sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i)$$

性质 3 的证明：

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

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

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

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

$$=Var(X)+Var(Y)$$

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

当 X,Y 相互独立时，

$$E[(X-E(X))(Y-E(Y))]$$

$$=[E(X-E(X))][E(Y-E(Y))]=0$$

$$\mathrm{Var}(X\pm Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$$

性质4 对任意常数C, $Var(X) \le E(X-C)^2$,

当且仅当$C = E(X)$时等号成立

性质5 若$Var(X)$存在，则$Var(X)=0$的充要条件是:

$$P(X=E(X))=1$$

例10 二项分布$X\sim B(n,p)$的方差

若设 $X_i=\begin{cases}1 & \text{如第}i\text{次试验成功}\\ 0 & \text{如第}i\text{次试验失败}\end{cases}$ $i=1,2,\ldots,n$

则 $X=\sum_{i=1}^{n}X_i$ 是n次试验中“成功”的次数

由于$X_1,X_2,\ldots,X_n$相互独立

于是 $\mathrm{Var}(X)=\sum_{i=1}^{n}\mathrm{Var}(X_i)=np(1-p)$

标准化随机变量

设随机变量X的方差$Var(X)$存在，且$D(X)>0$令

$$X^* = \frac{X - E(X)}{\sqrt{\mathrm{Var}(X)}}$$

其中$E(X)$是X的数学期望，求$E(X^*)$和$\mathrm{Var}(X^*)$.


休息片刻

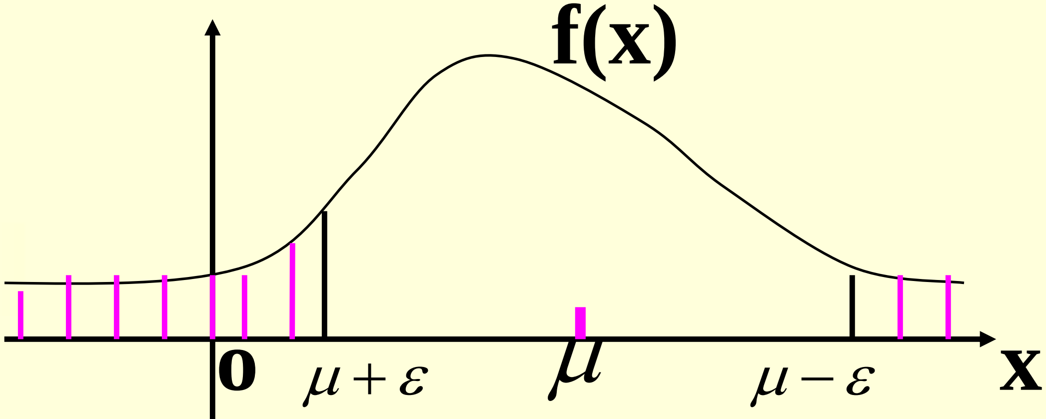
切比雪夫不等式

设随机变量X有期望 μ 和方差σ^2，则对于任给$\varepsilon>0$,

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

或

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



如取$\varepsilon = 3\sigma$

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

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

证明： 设*X~f(x)*

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

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

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

例11 已知正常男性成人血液中，每一毫升白细胞数平均是7300，标准差是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，$E(X)=7300$, $Var(X)=700^2$

所求为 $P(5200 \le X \le 9400)$

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

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

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

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

由切比雪夫不等式

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

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

即估计每毫升白细胞数在5200~9400之间的概率不小于8/9 .

例12 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?

解： 设X为n次试验中，事件A出现的次数，

则 $X \sim B(n, 0.75)$

$$E(X)=0.75n,\ Var(X)=0.75*0.25n=0.1875n$$

所求为满足

$$P\left(0.74<\frac{X}{n}<0.76\right)\geq 0.90$$

的最小的n.

$P(0.74 < \frac{X}{n} < 0.76)$ 可改写为

$P(0.74n < X < 0.76n)$

$= P(-0.01n < X - 0.75n < 0.01n)$

$= P\{|X - E(X)| < 0.01n\}$

在切比雪夫不等式中取 $\varepsilon = 0.01n$，则

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

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

依题意，取 $1-\dfrac{1875}{\boldsymbol{n}} \geq 0.9$

解得

$$n \geq \frac{1875}{1-0.9} = 18750$$

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

第四章 作业2：

10,13,14,17,20,24,34,35