西安電子科技大學
XIDIAN UNIVERSITY

第4章 随机变量的数字特征

概率论与数理统计课程组

CHAPTER 4

随机变量的数字特征

能描述随机变量某些特征的常数

CHAPTER 4

随机变量的数字特征

4.1 数学期望

除了分布函数和分布律/概率密度，还可以怎样表示随机试验的结果和特点？

在一些实际问题中，我们除分布函数外，更关心的是随机变量的某些特征

例如

- 在评定某地区粮食产量水平时，最关心的是平均产量
- 检查一批棉花质量时，既要注意纤维平均长度，又要注意纤维长度与平均长度的偏离程度
- 考察西安市居民的家庭收入情况，既要知家庭的年平均收入，又要研究贫富的差异程度

离散型随机变量的数学期望

定义 设离散型随机变量X的分布律为

$$P(X=x_k)=p_k \ (k = 1, 2, 3, \ldots)$$

若级数 $\sum_k x_k p_k$ 绝对收敛，则称 $\sum_k x_k p_k$ 的和为随机变量X的**数学期望**

记为$E(X)$或EX，即 $E(X)=\sum_k x_k p_k$

数学期望简称**期望**，又称为**均值**

关于定义的几点说明

1° $E(X)$是一个实数,而非变量,它是一种**加权平均**,与一般的平均值不同，它从本质上体现了随机变量 X 取可能值的**真正的平均值**, 也称均值.

2° **级数的绝对收敛性**保证了级数的和不随级数各项次序的改变而改变，之所以这样要求是因为数学期望是反映随机变量X 取可能值的平均值,它不应随可能值的排列次序而改变.

3° 随机变量的数学期望与一般变量的算术平均值不同.

连续型随机变量的数学期望

定义

设连续型随机变量X的概率密度为$f(x)$，

若积分$\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，即$\int_{-\infty}^{+\infty} |x| f(x)dx < +\infty$

则称$\int_{-\infty}^{+\infty} xf(x)dx$ 的值为随机变量X的**数学期望**

记为 $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$

数学期望$E(X)$完全由随机变量X的概率分布决定

例 设随机变量 X 服从参数为 p 的 0-1 分布，即 $X \sim B(1, p)$，求 $E(X)$

解 X 的分布律为 $P(X=0)=1-p$，$P(X=1)=p$

X 的数学期望为 $E(X)=0\times(1-p)+1\times p=p$

例 设随机变量 X 服从参数为 n、p 的二项分布，即 $X \sim B(n, p)$，求 $E(X)$

解 X 的分布律为 $P(X=k)=C_n^k p^k q^{n-k}$，$k=0,1,2,\cdots,n$，$q=1-p$

X 的数学期望为 $E(X)=\sum\limits_{k=0}^{n} kC_n^k p^k q^{n-k}=\sum\limits_{k=0}^{n} k\dfrac{n!}{k!(n-k)!}p^k q^{n-k}$

$$=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}q^{(n-1)-(k-1)}=np\sum_{k=1}^{n}C_{n-1}^{k-1}p^{k-1}q^{(n-1)-(k-1)}=np(p+q)^{n-1}=np$$

例 设随机变量 X 服从参数为 p 的几何分布，求 $E(X)$

解 X 的分布律为 $P(X=k)=q^{k-1}p,\quad k=1,2,\cdots,\quad 0<p<1, q=1-p$

X 的数学期望为 $E(X)=\sum_{k=1}^{\infty}k\cdot q^{k-1}p = p\left(1+2q+3q^2+\cdots\right)=p\left(q+q^2+q^3\cdots\right)'$

$$=p\left(\frac{q}{1-q}\right)'=\frac{p}{(1-q)^2}=\frac{1}{p}$$

例 设对于服从参数为 N、M、n 的超几何分布随机变量 X，求 $E(X)$

解 X 的分布律为 $P(X=k)=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n},\quad k=0,1,2,\cdots,\min\{M,n\}$

X 的数学期望为 $E(X)=\dfrac{nM}{N}$

例 设$X \sim P(\lambda)$，求 $E(X)$

解 X 的分布律为 $P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!} \quad k=0,1,\cdots \quad \lambda>0$

X 的数学期望为 $E(X)=\sum\limits_{k=0}^{\infty} k\dfrac{\lambda^k e^{-\lambda}}{k!}=\lambda e^{-\lambda}\sum\limits_{k=1}^{\infty}\dfrac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda$

例 设$X \sim U(a,b)$，求 $E(X)$

解 X 的概率密度为 $f(x)=\begin{cases}\dfrac{1}{b-a} & a<x<b \\ 0 & \text{其他}\end{cases}$

X 的数学期望为 $E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_a^b \dfrac{x}{b-a}dx=\dfrac{a+b}{2}$

即均匀分布的数学期望位于区间的中点

例 设$X \sim E(\lambda)$，求 $E(X)$

解 X 的概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$

X 的数学期望为 $E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_0^{+\infty} x\cdot\lambda e^{-\lambda x}dx=-\int_0^{+\infty} xde^{-\lambda x}=-xe^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty} e^{-\lambda x}dx=\dfrac{1}{\lambda}$

例 设$X \sim N(\mu, \sigma^2)$，求 $E(X)$

解 X 的概率密度为 $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，　$-\infty<x<+\infty$

X 的数学期望为 $E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_{-\infty}^{+\infty} x\cdot\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$

令$t=(x-\mu)/\sigma$

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

即参数μ是数学期望

例 设有10个同种电子元件，其中2个废品。装配仪器时，从这10个中任取1个，若是废品，扔掉后重取1只，求在取到正品之前已取出的废品数X 的期望。

解 X 的分布律为：

X	0	1	2
p_k	$\frac{8}{10}$	$\frac{2}{10}\times\frac{8}{9}$	$\frac{2}{10}\times\frac{1}{9}$

X	0	1	2
p_k	$\frac{4}{5}$	$\frac{8}{45}$	$\frac{1}{45}$

$$E(X)=0\times\frac{4}{5}+1\times\frac{8}{45}+2\times\frac{1}{45}=\frac{2}{9}$$

例 设一台机器一天内发生故障的概率为0.2，机器发生故障时全天停工。若一周5个工作日里无故障，可获利10万元；发生一次故障获利5万元；发生2次故障获利0元，发生3次或以上故障亏损2万元，求一周内期望利润是多少？

解 设X 表示一周5天内机器发生故障天数，则 $X\sim B(5,0.2)$

设Y 表示一周内所获利润，则 $P(Y=10)=P(X=0)=(1-0.2)^5=0.328$

其余同理可得，则Y 的分布律为

Y	-2	0	5	10
p_k	0.057	0.205	0.410	0.328

$E(Y)=5.216$

例 有2个相互独立工作的电子装置，它们的寿命X_1、X_2服从同一指数分布，概率密度为：

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}\quad \lambda>0$$

若将这2个电子装置串联联接组成整机，求整机寿命N（以小时计）的数学期望。

解 X_1、X_2分布函数为 $F(x)=\begin{cases}1-e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$

串联情况下，$N=\min\{X_1、X_2\}$ 分布函数为

$$F_{\min}(x)=1-[1-F(x)]^2=\begin{cases}1-e^{-2\lambda x} & x>0\\ 0 & x\le 0\end{cases}\quad\Rightarrow f_{\min}(x)=\begin{cases}2\lambda e^{-2\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

$$E(X_1)=\int_0^{+\infty}\lambda xe^{-\lambda x}dx=-xe^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-\lambda x}dx=-\frac{1}{\lambda}e^{-\lambda x}\Big|_0^{+\infty}=\frac{1}{\lambda}$$

因此 $E(N)=\frac{1}{2\lambda}$ 并联？

只要求得一般指数分布的期望，即可同理得$E(N)$

定理 不必用Y的分布律或概率密度求$E(Y)$

设Y是随机变量X的函数:$Y=g(X)$，g是连续函数

(i) 如果X是离散型随机变量，它的分布律是$P(X=x_k)=p_k$ $(k = 1, 2, 3, \ldots)$

若级数$\sum_k g(x_k)p_k$绝对收敛，则有

$$E(Y)=E\left[g(X)\right]=\sum_k g(x_k)p_k$$

(ii) 如果X是连续型随机变量，它的概率密度是 $f(x)$，

若级数$\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛，则有

$$E(Y)=E\left[g(X)\right]=\int_{-\infty}^{+\infty} g(x)f(x)dx$$

推广到两个随机变量的函数的情况

设Z是随机变量X、Y的函数:$Z=g(X, Y)$，g是连续函数，Z是一维随机变量

若离散型随机变量(X, Y) 分布律是$P(X=x_i, Y=y_j)=p_{ij}$ $(i, j = 1, 2, 3, \ldots)$

$$E(Z)=E\left[g(X,Y)\right]=\sum_j\sum_i g(x_i,y_j)p_{ij}$$

设上式的级数绝对收敛

若二维随机变量(X, Y)的概率密度是 $f(x, y)$，则有

$$E(Z)=E\left[g(X,Y)\right]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$$

设上式的积分绝对收敛

例 已知某零件的横截面是个圆，对横截面的直径X进行测量，其值在区间(1, 2)上均匀分布，求横截面面积S的数学期望。

解 X概率密度为 $f(x)=\begin{cases}1, & 1<x<2\\ 0, & \text{其他}\end{cases}$ $S=\dfrac{\pi X^2}{4}$

$$E(S)=\int_{-\infty}^{+\infty}\frac{\pi}{4}x^2 f(x)dx=\int_1^2\frac{\pi x^2}{4}dx=\frac{7\pi}{12}$$

例 设二维随机变量(X, Y)联合分布律如右图，求随机变量 $Z=\sin\dfrac{\pi(X+Y)}{2}$ 的数学期望。

解
$$E(Z)=E[\sin\frac{\pi(X+Y)}{2}]$$
$$=\sin\frac{\pi(0+0)}{2}\times 0.1+\sin\frac{\pi(1+0)}{2}\times 0.15+\sin\frac{\pi(0+1)}{2}\times 0.25$$
$$+\sin\frac{\pi(1+1)}{2}\times 0.2+\sin\frac{\pi(0+2)}{2}\times 0.15+\sin\frac{\pi(1+2)}{2}\times 0.15$$
$$=0.25$$

$X \backslash Y$	0	1	2
0	0.1	0.25	0.15
1	0.15	0.2	0.15

例 设 (X, Y) 的分布律为

$Y \backslash X$	1	2	3
-1	0.2	0.1	0
0	0.1	0	0.3
1	0.1	0.1	0.1

求: $E(X)$, $E(Y)$, $E(Y/X)$, $E[(X-Y)^2]$.

解 X 的分布律为

X	1	2	3
p	0.4	0.2	0.4

得 $E(X)=1\times0.4+2\times0.2+3\times0.4=2.$

Y 的分布律为

Y	-1	0	1
p	0.3	0.4	0.3

得 $E(Y)=-1\times0.3+0\times0.4+1\times0.3=0.$

由于

p	0.2	0.1	0.1	0.1	0.1	0.3	0.1
(X,Y)	(1,−1)	(1,0)	(1,1)	(2,−1)	(2,1)	(3,0)	(3,1)
Y/X	-1	0	1	$-1/2$	$1/2$	0	$1/3$

于是

$$E\left(\frac{Y}{X}\right)=-1\times0.2+0\times0.1+1\times0.1-\frac{1}{2}\times0.1+\frac{1}{2}\times0.1+0\times0.3+\frac{1}{3}\times0.1=\frac{1}{15}.$$

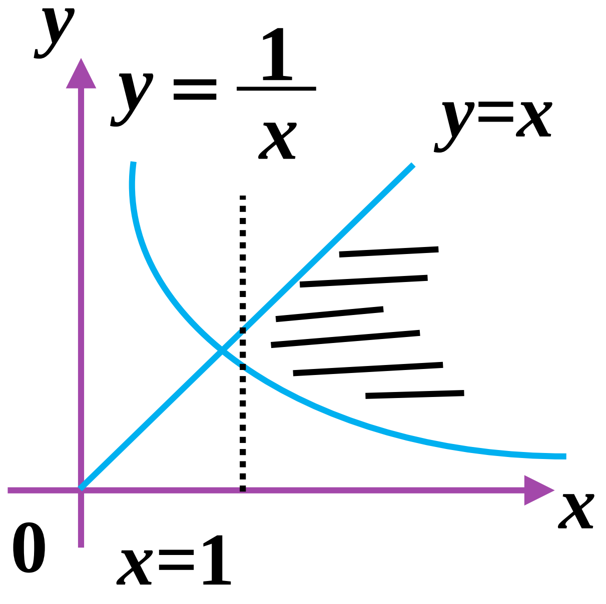
p	0.2	0.1	0.1	0.1	0.1	0.3	0.1
(X,Y)	(1,–1)	(1,0)	(1,1)	(2,–1)	(2,1)	(3,0)	(3,1)
$(X-Y)^2$	4	1	0	9	1	9	4

得 $E[(X-Y)^2]=4\times0.3+1\times0.2+0\times0.1+9\times0.4=5.$

例 设二维随机变量(X, Y)概率密度为

$$f(x,y)=\begin{cases}\dfrac{3}{2x^3y^2} & \dfrac{1}{x}<y<x,x>1\\ 0 & 其他\end{cases}$$

求数学期望 $E(Y),E(\dfrac{1}{XY})$



解 $E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dydx=\int_{1}^{+\infty}\int_{\frac{1}{x}}^{x}\dfrac{3}{2x^3y}dydx=\dfrac{3}{2}\int_{1}^{+\infty}\dfrac{1}{x^3}\ln y|_{\frac{1}{x}}^{x}dx$

$=3\int_{1}^{+\infty}\dfrac{\ln x}{x^3}dx=-\dfrac{3}{2}\dfrac{\ln x}{x^2}|_{1}^{+\infty}+\dfrac{3}{2}\int_{1}^{+\infty}\dfrac{1}{x^3}dx=\dfrac{3}{4}$ 也可以先求 $E(Y)=\int_{-\infty}^{+\infty}yf_Y(y)dy$

$E(\dfrac{1}{XY})=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\dfrac{1}{xy}f(x,y)dydx=\int_{1}^{+\infty}dx\int_{\frac{1}{x}}^{x}\dfrac{3}{2x^4y^3}dy=\int_{1}^{+\infty}\dfrac{3}{2x^4}[-\dfrac{1}{2y^2}]|_{\frac{1}{x}}^{x}dx$

$=\dfrac{3}{4}\int_{1}^{+\infty}(-\dfrac{1}{x^6}+\dfrac{1}{x^2})dx=\dfrac{3}{4}(-\dfrac{1}{5}+1)=\dfrac{3}{5}$

例 某商店经销某种商品，每周进货量X与需求量Y是相互独立的随机变量，且都在区间[10, 20]上均匀分布。商店每售出一单位商品可获利1000元；若需求量超过进货量，商店可从其他处调剂供应，这时每单位商品可获利500元；试计算此商店经销该种商品每周所获利润的数学期望。

解 设Z表示该种商品每周所得的利润，则

$$Z = g(X,Y) = \begin{cases} 1000Y, & \text{若} Y \le X \\ 500(X+Y), & \text{若} Y > X \end{cases}$$

X和Y相互独立，因此(X, Y)联合概率密度为 $f(x,y) = \begin{cases} 1/100, & 10 \le x \le 20, 10 \le y \le 20 \\ 0, & \text{其他} \end{cases}$

$$E(Z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f(x,y)dxdy$$

$$= \int_{10}^{20} dx \int_{10}^{x} 1000y \times 1/100\, dy + \int_{10}^{20} dx \int_{x}^{20} 500(x+y) \times 1/100\, dy \approx 14166.7$$

数学期望的性质

1° 设C是常数，则有$E(C)=C$

2° 设X是一个随机变量，C是常数，则有
$E(CX)=CE(X)$

3° 设X、Y是两个随机变量，则有
$E(X+Y)=E(X)+E(Y)$

推广到有限个随机变量之和，有
$E(aX+bY+c)=aE(X)+bE(Y)+c$

4° 设X、Y是两个**相互独立**的随机变量，则有
$E(XY)=E(X)E(Y)$

可推广到有限个相互独立的随机变量之积

证明

1° C是常数，$P(X=C)=1$，$E(X)=E(C)=1\times C=C$

下面仅对连续型随机变量给予证明

$$E(CX)=\int_{-\infty}^{+\infty}Cxf(x)dx=C\int_{-\infty}^{+\infty}xf(x)dx=CE(X)$$

$$E(X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dxdy$$

$$=E(X)+E(Y)$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_X(x)f_Y(y)dxdy=\int_{-\infty}^{+\infty}xf_X(x)dx\int_{-\infty}^{+\infty}yf_Y(y)dy$$

$$=E(X)E(Y)$$

例 设随机变量X_1, X_2, X_3, X_4 相互独立，$X_i \sim U(0, 2i)$，求以下行列式的数学期望$E(Y)$。

$$Y = \begin{vmatrix} X_1 & X_2 \\ X_3 & X_4 \end{vmatrix}$$

解 $E(X_i) = i, \quad i = 1,2,3,4$

$Y = X_1X_4 - X_2X_3$

由条件 $E(Y) = E(X_1X_4) - E(X_2X_3)$

$= E(X_1)E(X_4) - E(X_2)E(X_3)$

$= 1\times4 - 2\times3 = -2$

例 一民航送客车载有20位旅客自机场出发，旅客有10个车站可以下车，如到达一个车站没有旅客下车就不停车，以X 表示停车的次数，求 $E(X)$。

（设每位旅客在各个车站下车是等可能的，并设各旅客是否下车相互独立）

解 引入随机变量

$$X_i = \begin{cases} 0 & 第i站没有人下车 \\ 1 & 第i站有人下车 \end{cases} \quad i = 1, 2, \cdots, 10$$

易知 $X = X_1 + X_2 + \cdots + X_{10}$

$$E(X_i) = P(X_i = 1) = P(第i站有人下车) = 1 - (\frac{9}{10})^{20}$$

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_{10}) = 10\left[1 - (\frac{9}{10})^{20}\right] = 8.784$$

本题将X分解成数个随机变量之和，然后利用随机变量和的数学期望等于随机变量数学期望之和求数学期望，这种方法具有一定普遍意义

本节回顾

离散型数学期望的定义及计算

设离散型随机变量X的分布律为 $P(X=x_k)=p_k \ (k=1,2,3,\dots)$

若级数 $\sum_k x_k p_k$ 绝对收敛，则称 $\sum_k x_k p_k$ 的和为随机变量X的**数学期望** $E(X)=\sum_k x_k p_k$

连续型数学期望的定义及计算

设连续型随机变量X的概率密度为$f(x)$，若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，即 $\int_{-\infty}^{+\infty} |x| f(x)dx < +\infty$

则称 $\int_{-\infty}^{+\infty} xf(x)dx$ 的值为随机变量X的**数学期望** 记为 $E(X)=\int_{-\infty}^{+\infty} xf(x)dx$

CHAPTER 4

随机变量的数字特征

4.2 方差

引例

- 有一批器件寿命为：一半约950小时，另一半约1050小时→平均寿命为1000小时；
  另一批器件寿命为：　一半约1300小时，另一半约700小时→平均寿命为1000小时。
  问题：哪批器件的质量更好？

单从平均寿命这一指标无法判断，需进一步考察寿命X与均值1000小时的偏离程度，方差正是体现这种意义的数学特征

例如　$E\{|X-E(X)|\}$　　$E\{[X-E(X)]^2\}$

定义 方差本质是X的函数$g(X)=[X-E(X)]^2$的期望

设X是一个随机变量，若 $E\left\{[X-E(X)]^2\right\}$ 存在，则称 $E\left\{[X-E(X)]^2\right\}$ 为X的**方差**，记为 $D(X)$或DX或$Var(X)$，即

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

在应用上还引入量 $\sqrt{D(X)}$，记为$\sigma(X)$，称为**标准差或均方差**，它是与随机变量X具有相同量纲的量

方差$D(X)$刻画了X取值的分散程度，它是衡量X取值分散程度的一个尺度：若X取值比较集中，则$D(X)$较小，反之，若X取值比较分散，则$D(X)$较大

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$P(X=x_k)=p_k$ $(k = 1, 2, 3, \ldots)$ 则X的**方差**为

$$D(X)=\sum_k\left[x_k-E(X)\right]^2 p_k$$

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$f(x)$，则X的方差为

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

此外，利用数学期望的性质，可得方差得计算公式，有

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

证

$$\begin{aligned}D(X)&=E\left\{\left[X-E(X)\right]^2\right\}\\&=E\left\{X^2-2XE(X)+\left[E(X)\right]^2\right\}\\&=E(X^2)-2E(X)E(X)+\left[E(X)\right]^2\\&=E(X^2)-\left[E(X)\right]^2\end{aligned}$$

例 设X 服从0-1分布，分布律为$P(X=0)=1-p$, $P(X=1)=p$, 求 $D(X)$

解

$$E(X)=0\cdot(1-p)+1\cdot p=p$$

$$E(X^2)=0^2\cdot(1-p)+1^2\cdot p=p$$

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

例 设$X\sim P(\lambda)$，求 $D(X)$

解

X 的分布律为 $P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}\quad k=0,1,\cdots\quad \lambda>0$

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

$$=\sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k e^{-\lambda}}{k!}+\lambda=\lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda=\lambda^2 e^{-\lambda}e^{\lambda}+\lambda=\lambda^2+\lambda$$

X 的方差为 $D(X)=E(X^2)-[E(X)]^2=\lambda$

泊松分布期望与方差相等

例 设$X \sim U(a, b)$，求 $D(X)$

解 X 的概率密度为

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

之前算得X 的数学期望为 $E(X)=\dfrac{a+b}{2}$

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

$$=\frac{b^3-a^3}{3(b-a)}=\frac{a^2+b^2+ab}{3}$$

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

$$=\frac{a^2+b^2+ab}{3}-\frac{a^2+b^2+2ab}{4}=\frac{(b-a)^2}{12}$$

例 设X 服从指数分布，概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$

其中$\lambda>0$，求 $E(X)$、$D(X)$

解 $E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^{+\infty}\lambda xe^{-\lambda x}dx$

$$=-xe^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-\lambda x}dx=\frac{1}{\lambda}$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2 f(x)dx=\int_0^{+\infty}x^2\lambda e^{-\lambda x}dx$$

$$=-x^2e^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty}2xe^{-\lambda x}dx=\frac{2}{\lambda^2}$$

于是 $D(X)=E(X^2)-[E(X)]^2=\dfrac{2}{\lambda^2}-\dfrac{1}{\lambda^2}=\dfrac{1}{\lambda^2}$

即指数分布的期望恰为参数λ的倒数，方差是期望的平方

方差的性质

1° 设C是常数，则有$D(C)=0$

2° 设X是一个随机变量， C是常数，则有$D(X+C)=D(X)$

3° 设X是一个随机变量， C是常数，则有$D(CX)=C^2D(X)$

4° 设X、Y是两个相互独立的随机变量，则有 $D(X+Y)=D(X)+D(Y)$

推广到有限个相互独立随机变量之和，有
$D(aX+bY+c)=a^2D(X)+b^2D(Y)$

一般地，若X、Y是任意两随机变量（不要求相互独立），有

$$D(X+Y)=D(X)+D(Y)+2E\left\{[X-E(X)][Y-E(Y)]\right\}$$

5° $D(X)=0$的充要条件是X以概率1取常数$E(X)$，即 $P(X=E(X))=1$

证明

1° C是常数， $D(C)=E\left\{[C-E(C)]^2\right\}=0$

2° $D(X+C)=E\left\{[X+C-E(X+C)]^2\right\}$

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

3° $D(CX)=E\left[(CX)^2\right]-[E(CX)]^2$

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

$$=C^2\left\{E(X^2)-[E(X)]^2\right\}=C^2D(X)$$

4° $D(X+Y)=E\left\{[(X+Y)-E(X+Y)]^2\right\}$

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

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

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

$$=D(X)+D(Y)+2E\left\{[X-E(X)][Y-E(Y)]\right\}$$

$$2E\left\{[X-E(X)][Y-E(Y)]\right\}$$

$$=2E\left\{XY-XE(Y)-YE(X)+E(X)E(Y)\right\}$$

$$=2\left\{E(XY)-E(X)E(Y)-E(X)E(Y)+E(X)E(Y)\right\}$$

$$=2\left\{E(XY)-E(X)E(Y)\right\}$$ 协方差

故 $D(X+Y)=D(X)+D(Y)+2E\{E(XY)-E(X)E(Y)\}$

若X、Y **相互独立**，则有 $D(X+Y)=D(X)+D(Y)$

5° 充分性

设$P\{X=E(X)\}=1$，则有 $P\left\{X^2=[E(X)]^2\right\}=1$

于是 $D(X)=E(X^2)-[E(X)]^2=0$

必要性证明见第5章*Chebyshev*不等式处

例 设$X\sim B(n,p)$，求 $E(X)$、$D(X)$

解 随机变量X是事件A在n重伯努利试验中发生的次数，$(X=0,1,2,\dots,n)$，每次事件相互独立，$P(A)=p$，记

$$X_k=\begin{cases}1 & A\text{在第}k\text{次试验发生}\\ 0 & A\text{在第}k\text{次试验不发生}\end{cases}\qquad k=1,2,\cdots n$$

于是$X_1,X_2,\dots,X_n$相互独立，服从同一分布即0-1分布

X_k	0	1
p_k	$1-p$	p

$$X=X_1+X_2+\cdots X_n$$

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

定理

以n,p为参数的二项分布变量，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

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

解 先求标准正态变量Z的数学期望和方差 $Z=\dfrac{X-\mu}{\sigma}$

Z的概率密度为 $\varphi(z)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$

于是 $E(Z)=\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z e^{-\frac{z^2}{2}}dz=0$

$$D(Z)=E(Z^2)=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}z^2e^{-\frac{z^2}{2}}dz=-\frac{1}{\sqrt{2\pi}}ze^{-\frac{z^2}{2}}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=1$$

因为 $X=\mu+\sigma Z$

$$E(X)=E(\mu+\sigma Z)=\mu,\quad D(X)=D(\mu+\sigma Z)=\sigma^2D(Z)=\sigma^2$$

即正态分布的两个参数 μ, σ^2分别是其期望和方差

随机变量线性变换时的数学期望和方差

设随机变量X具有数学期望$E(X)=\mu$，方差$D(X)=\sigma^2\neq 0$，

记 $X^*=\dfrac{X-\mu}{\sigma}$ 则$E(X^*)=0,\ D(X^*)=1$

证

$$E(X^*)=\frac{1}{\sigma}E(X-\mu)=\frac{1}{\sigma}\left[E(X)-\mu\right]=0$$

$$D(X^*)=E(X^{*2})-\left[E(X^*)\right]^2$$

不一定是$N(0, 1)$

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

$$=\frac{1}{\sigma^2}E\left[(X-\mu)^2\right]=\frac{\sigma^2}{\sigma^2}=1$$

X^*称为X的标准化变量

定理

独立的n个正态变量的线性组合仍服从正态分布

若 $X_i \sim N(\mu_i, \sigma_i^2)$ $i=1,2,\cdots n$ 且它们**相互独立**

它们的线性组合也服从正态分布，即有

$$C_0 + C_1X_1 + C_2X_2 + \cdots + C_nX_n \sim N(C_0 + C_1\mu_1 + \cdots + C_n\mu_n, C_1^2\sigma_1^2 + C_2^2\sigma_2^2 + \cdots + C_n^2\sigma_n^2)$$

其中C_1、C_2、…、C_n为不全为0的常数

如： $X\sim N(1, 3)$, $Y\sim N(2, 4)$ 且X、Y相互独立，

则 $Z=2X-3Y \sim N(-4, 48)$

$D(Z)=2^2D(X)-3^2D(Y)$，对吗？

在多个变量和的方差计算中，负号不会造成**方差**相减，
不同变量在加减时它们的方差都是**相加**的，因为$(-1)^2=1$

例 设活塞的直径(以cm计) 和汽缸的直径X,Y相互独立，$X\sim N(22.40, 0.03^2)$，$Y\sim N(22.50, 0.04^2)$，任取一只活塞，任取一只汽缸，求活塞能装入汽缸的概率。

解 依照题意，需求 $P(X<Y)=P(X-Y<0)$

由于 $X-Y\sim N(-0.10, 0.05^2)$

故有 $P(X<Y)=P(X-Y<0)$

$$=\Phi\left(\frac{0-(-0.10)}{0.05}\right)$$

$$=\Phi(2)=0.9772$$

例　设连续型随机变量X的概率密度为 $f(x)=\begin{cases}\cos x, & 0\le x\le \dfrac{\pi}{2},\\ 0, & \text{其它}.\end{cases}$

求随机变量 $Y=X^2$ 的方差 $D(Y)$.

解　$E(X^2)=\int_{-\infty}^{\infty}x^2 f(x)\mathrm{d}x=\int_0^{\frac{\pi}{2}}x^2\cos x\,\mathrm{d}x=\dfrac{\pi^2}{4}-2,$

$E(X^4)=\int_{-\infty}^{\infty}x^4 f(x)\mathrm{d}x=\int_0^{\frac{\pi}{2}}x^4\cos x\,\mathrm{d}x=\dfrac{\pi^4}{16}-3\pi^2+24,$

因为 $D(X)=E(X^2)-[E(X)]^2,$

所以 $D(X^2)=E(X^4)-[E(X^2)]^2=\dfrac{\pi^4}{16}-3\pi^2+24-\left(\dfrac{\pi^2}{4}-2\right)^2=20-2\pi^2.$

例 设 $X \sim \begin{pmatrix} -2 & 0 & 1 & 3 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{12} & \frac{1}{12} \end{pmatrix}$ 求 $D(2X^3+5)$.

解 $D(2X^3+5)=D(2X^3)+D(5)=4D(X^3)=4[E(X^6)-(E(X^3))^2]$

$$E(X^6)=(-2)^6\times\frac{1}{3}+0^6\times\frac{1}{2}+1^6\times\frac{1}{12}+3^6\times\frac{1}{12}=\frac{493}{6},$$

$$[E(X^3)]^2=\left[(-2)^3\times\frac{1}{3}+0^3\times\frac{1}{2}+1^3\times\frac{1}{12}+3^3\times\frac{1}{12}\right]^2=\frac{1}{9},$$

故 $D(2X^3+5)=4[E(X^6)-(E(X^3))^2]=\frac{2954}{9}$.

常见离散型随机变量的数学期望与方差

分布名称	参数	分布律	数学期望	方差
0−1分布	$0<p<1$	$P(X=k)=p^k(1-p)^{1-k}\qquad k=0,1$	p	$p(1-p)$
二项分布 $B(n,p)$	$n\geq 1,\ 0<p<1$	$P(X=k)=C_n^k p^k(1-p)^{1-k}\qquad k=0,1,\ldots,n$	np	$np(1-p)$
泊松分布 $P(\lambda)$	$\lambda>0$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}\qquad k=0,1,\ldots$	λ	λ
几何分布	$0<p<1$	$P(X=k)=(1-p)^{k-1}p\qquad k=1,2,\ldots$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
超几何分布	$M<N,\ n<N$	$P(X=k)=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}\qquad k=0,1,2,\ldots,\min\{M,n\}$	$\dfrac{nM}{N}$	$\dfrac{nM(N-M)(N-n)}{N^2(N-1)}$

常见连续型随机变量的数学期望与方差

分布名称	参数	概率密度	数学期望	方差
均匀分布 $U(a, b)$	$a<b$	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b \\ 0, & 其他\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
指数分布 参数 λ	$\lambda>0$	$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0 \\ 0 & x\le 0\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
正态分布 $N(\mu, \sigma^2)$	$\sigma>0$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $(-\infty<x<+\infty)$	μ	σ^2

本节回顾

方差的定义及计算

设离散型随机变量X的分布律为$P(X=x_k)=p_k$, (k = 1, 2, 3, …)　　则X的方差为

$$D(X)=\sum_k [x_k-E(X)]^2 p_k$$

设连续型随机变量X的概率密度为$f(x)$，则X的方差为 $D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx$

利用数学期望的性质，可得方差得计算公式，有 $D(X)=E(X^2)-[E(X)]^2$