第四章 随机变量的数字特征

上一章介绍了随机变量的分布函数、概率密度和分布律，它们都能完整地描述随机变量，但在某些实际或理论问题中，人们感兴趣于某些能描述随机变量某一种特征的常数，例如，一篮球队上场比赛的运动员的身高是一个随机变量，人们常关心上场运动员的平均身高. 一个城市一户家庭拥有汽车的辆数是一个随机变量，在考察城市的交通情况时，人们关心户均拥有汽车的辆数. 评价棉花的质量时，既需要注意纤维的平均长度，又需要注意纤维长度与平均长度的偏离程度，平均长度较大，偏离程度较小，质量就较好. 这种由随机变量的分布所确定的，能刻画随机变量某一方面的特征的常数统称为数字特征，它在理论和实际应用中都很重要. 本章将介绍几个重要的数字特征：数学期望、方差、相关系数和矩.

§1 数学期望

先看一个例子. 一射手进行打靶练习，规定射入区域 e_2（图4－1）得 2 分，射入区域 e_1 得 1 分，脱靶，即射入区域 e_0，得 0 分. 射手一次射击得分数 X 是一个随机变量. 设 X 的分布律为

$$P\{X=k\}=p_k, k=0,1,2.$$

现在射击 N 次，其中得 0 分的有 a_0 次，得 1 分的有 a_1 次，得 2 分的有 a_2 次，$a_0+a_1+a_2=N$. 他射击 N 次得分的总和为 $a_0\times 0+a_1\times 1+a_2\times 2$. 于是平均一次射击的得分数为

$$\frac{a_0\times 0+a_1\times 1+a_2\times 2}{N}=\sum_{k=0}^{2}k\frac{a_k}{N}.$$

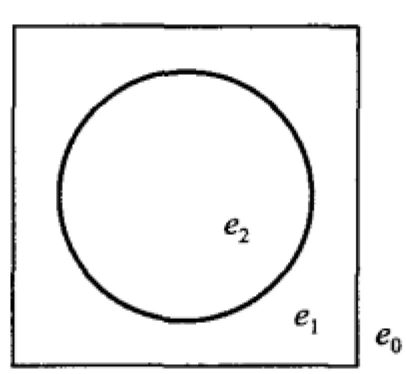


图 4－1

这里，a_k/N 是事件$\{X=k\}$的频率. 在第五章将会讲到，当 N 很大时，a_k/N 在一定意义下接近于事件$\{X=k\}$的概率 p_k. 就是说，在试验次数很大时，随机变量 X 的观察值的算术平均 $\sum_{k=0}^{2}ka_k/N$ 在一定意义下接近于 $\sum_{k=0}^{2}kp_k$. 我们称 $\sum_{k=0}^{2}kp_k$ 为随机变量 X 的数学期望或均值. 一般，有以下的定义.

定义 设离散型随机变量 X 的分布律为

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots.$$

若级数

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

绝对收敛,则称级数 $\sum\limits_{k=1}^{\infty} x_k p_k$ 的和为随机变量 X 的**数学期望**,记为 $E(X)$. 即

$$E(X) = \sum_{k=1}^{\infty} x_k p_k. \tag{1.1}$$

设连续型随机变量 X 的概率密度为 $f(x)$,若积分

$$\int_{-\infty}^{\infty} x f(x) \mathrm{d}x$$

绝对收敛,则称积分 $\int_{-\infty}^{\infty} x f(x) \mathrm{d}x$ 的值为随机变量 X 的**数学期望**,记为 $E(X)$. 即

$$E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x. \tag{1.2}$$

数学期望简称**期望**,又称为**均值**.

数学期望 $E(X)$ 完全由随机变量 X 的概率分布所确定. 若 X 服从某一分布,也称 $E(X)$ 是这一分布的数学期望.

例 1 某医院当新生儿诞生时,医生要根据婴儿的皮肤颜色、肌肉弹性、反应的敏感性、心脏的搏动等方面的情况进行评分,新生儿的得分 X 是一个随机变量. 据以往的资料表明 X 的分布律为

X	0	1	2	3	4	5	6	7	8	9	10
p_k	0.002	0.001	0.002	0.005	0.02	0.04	0.18	0.37	0.25	0.12	0.01

试求 X 的数学期望 $E(X)$.

解
$$\begin{aligned} E(X) &= 0\times0.002+1\times0.001+2\times0.002+3\times0.005+4\times0.02 \\ &\quad +5\times0.04+6\times0.18+7\times0.37+8\times0.25+9\times0.12+10\times0.01 \\ &= 7.15(\text{分}) \end{aligned}$$

这意味着,若考察医院出生的很多新生儿,例如 1000 个,那么一个新生儿的平均得分约 7.15 分,1000 个新生儿共得分约 7150 分. □

例 2 有两个相互独立工作的电子装置,它们的寿命(以小时计)$X_k(k=1,2)$服从同一指数分布,其概率密度为

$$f(x)=\begin{cases} \dfrac{1}{\theta}\mathrm{e}^{-x/\theta}, & x>0, \\ 0, & x\leqslant 0, \end{cases} \quad \theta>0.$$

若将这两个电子装置串联连接组成整机,求整机寿命(以小时计)N 的数学期望.

解 $X_k(k=1,2)$的分布函数为

$$F(x)=\begin{cases}1-e^{-x/\theta}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

由第三章§5的(5.12)式 $N=\min\{X_1,X_2\}$的分布函数为

$$F_{\min}(x)=1-[1-F(x)]^2=\begin{cases}1-e^{-2x/\theta}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

因而 N 的概率密度为

$$f_{\min}(x)=\begin{cases}\dfrac{2}{\theta}e^{-2x/\theta}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

于是 N 的数学期望为

$$E(N)=\int_{-\infty}^{\infty}xf_{\min}(x)\mathrm{d}x=\int_0^{\infty}\frac{2x}{\theta}e^{-2x/\theta}\mathrm{d}x=\frac{\theta}{2}.$$ □

例 3 按规定，某车站每天 8:00～9:00，9:00～10:00 都恰有一辆客车到站，但到站的时刻是随机的，且两者到站的时间相互独立. 其规律为

到站时刻	8:10 9:10	8:30 9:30	8:50 9:50
概率	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

一旅客 8:20 到车站，求他候车时间的数学期望.

解 设旅客的候车时间为 X（以分计）. X 的分布律为

X	10	30	50	70	90
p_k	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}\times\frac{1}{6}$	$\frac{1}{6}\times\frac{3}{6}$	$\frac{1}{6}\times\frac{2}{6}$

在上表中，例如

$$P\{X=70\}=P(AB)=P(A)P(B)=\frac{1}{6}\times\frac{3}{6},$$

其中 A 为事件“第一班车在 8:10 到站”，B 为“第二班车在 9:30 到站”. 候车时间的数学期望为

$$\begin{aligned}E(X)&=10\times\frac{3}{6}+30\times\frac{2}{6}+50\times\frac{1}{36}+70\times\frac{3}{36}+90\times\frac{2}{36}\\&=27.22(\text{分}).\end{aligned}$$ □

例 4 某商店对某种家用电器的销售采用先使用后付款的方式. 记使用寿命为 X（以年计），规定：

$X\leqslant 1$，一台付款 1 500 元；

$1<X\leqslant 2$，一台付款 2 000 元；

$2<X\leqslant 3$,一台付款 2 500 元;

$X>3$,一台付款 3 000 元.

设寿命 X 服从指数分布,概率密度为

$$f(x)=\begin{cases}\dfrac{1}{10}\mathrm{e}^{-x/10}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

试求该商店一台这种家用电器收费 Y 的数学期望.

解 先求出寿命 X 落在各个时间区间的概率.即有

$$P\{X\leqslant 1\}=\int_0^1\frac{1}{10}\mathrm{e}^{-x/10}\mathrm{d}x=1-\mathrm{e}^{-0.1}=0.0952,$$

$$P\{1<X\leqslant 2\}=\int_1^2\frac{1}{10}\mathrm{e}^{-x/10}\mathrm{d}x=\mathrm{e}^{-0.1}-\mathrm{e}^{-0.2}=0.0861,$$

$$P\{2<X\leqslant 3\}=\int_2^3\frac{1}{10}\mathrm{e}^{-x/10}\mathrm{d}x=\mathrm{e}^{-0.2}-\mathrm{e}^{-0.3}=0.0779,$$

$$P\{X>3\}=\int_3^\infty\frac{1}{10}\mathrm{e}^{-x/10}\mathrm{d}x=\mathrm{e}^{-0.3}=0.7408.$$

一台家用电器收费 Y 的分布律为

Y	1 500	2 000	2 500	3 000
p_k	0.095 2	0.086 1	0.077 9	0.740 8

得 $E(Y)=2\,732.15$,即平均一台收费 2732.15 元. □

例 5 在一个人数很多的团体中普查某种疾病,为此要抽验 N 个人的血,可以用两种方法进行.(i)将每个人的血分别去验,这就需验 N 次.(ii)按 k 个人一组进行分组,把从 k 个人抽来的血混合在一起进行检验,如果这混合血液呈阴性反应,就说明 k 个人的血都呈阴性反应,这样,这 k 个人的血就只需验一次.若呈阳性,则再对这 k 个人的血液分别进行化验.这样,k 个人的血总共要化验 $k+1$ 次.假设每个人化验呈阳性的概率为 p,且这些人的试验反应是相互独立的.试说明当 p 较小时,选取适当的 k,按第二种方法可以减少化验的次数.并说明 k 取什么值时最适宜.

解 各人的血呈阴性反应的概率为 $q=1-p$.因而 k 个人的混合血呈阴性反应的概率为 q^k,k 个人的混合血呈阳性反应的概率为 $1-q^k$.

设以 k 个人为一组时,组内每人化验的次数为 X,则 X 是一个随机变量,其分布律为

X	$\dfrac{1}{k}$	$\dfrac{k+1}{k}$
p_k	q^k	$1-q^k$

X 的数学期望为

$$E(X)=\frac{1}{k}q^k+\left(1+\frac{1}{k}\right)(1-q^k)=1-q^k+\frac{1}{k}.$$

N 个人平均需化验的次数为

$$N\left(1-q^k+\frac{1}{k}\right).$$

由此可知，只要选择 k 使

$$1-q^k+\frac{1}{k}<1,$$

则 N 个人平均需化验的次数 $<N$. 当 p 固定时，我们选取 k 使得

$$L=1-q^k+\frac{1}{k}$$

小于 1 且取到最小值，这时就能得到最好的分组方法.

例如，$p=0.1$，则 $q=0.9$，当 $k=4$ 时，$L=1-q^k+\frac{1}{k}$ 取到最小值. 此时得到最好的分组方法. 若 $N=1000$，此时以 $k=4$ 分组，则按第二种方法平均只需化验

$$1000\left(1-0.9^4+\frac{1}{4}\right)=594(\text{次}).$$

这样平均来说，可以减少 40% 的工作量. □

例 6 设 $X\sim\pi(\lambda)$，求 $E(X)$.

解 X 的分布律为

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,2,\cdots,\quad \lambda>0.$$

X 的数学期望为

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

即 $E(X)=\lambda$. □

例 7 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

$$E(X)=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x=\int_a^b\frac{x}{b-a}\mathrm{d}x=\frac{a+b}{2}.$$

即数学期望位于区间 (a,b) 的中点. □

我们经常需要求随机变量的函数的数学期望，例如飞机机翼受到压力 $W=$

kV^2(V 是风速,$k>0$ 是常数)的作用,需要求 W 的数学期望,这里 W 是随机变量 V 的函数. 这时,可以通过下面的定理来求 W 的数学期望.

定理 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,它的分布律为 $P\{X=x_k\}=p_k, k=1,2,\cdots$,若 $\sum\limits_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛,则有

$$E(Y)=E[g(X)]=\sum_{k=1}^{\infty} g(x_k)p_k. \tag{1.3}$$

(ii) 如果 X 是连续型随机变量,它的概率密度为 $f(x)$,若 $\int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x$ 绝对收敛,则有

$$E(Y)=E[g(X)]=\int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x. \tag{1.4}$$

定理的重要意义在于当我们求 $E(Y)$时,不必算出 Y 的分布律或概率密度,而只需利用 X 的分布律或概率密度就可以了,定理的证明超出了本书的范围. 我们只对下述特殊情况加以证明.

证 设 X 是连续型随机变量,且 $y=g(x)$满足第二章§5 中定理的条件.

由第二章§5 中的(5.2)式知道随机变量 $Y=g(X)$的概率密度为

$$f_Y(y)=\begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta, \\ 0, & \text{其他}, \end{cases}$$

于是

$$E(Y)=\int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y=\int_{\alpha}^{\beta} y f_X[h(y)]\,|h'(y)|\,\mathrm{d}y.$$

当 $h'(y)$恒>0 时

$$E(Y)=\int_{\alpha}^{\beta} y f_X[h(y)]h'(y)\mathrm{d}y=\int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x.$$

当 $h'(y)$恒<0 时

$$\begin{aligned} E(Y) &= -\int_{\alpha}^{\beta} y f_X[h(y)]h'(y)\mathrm{d}y \\ &= -\int_{\infty}^{-\infty} g(x)f(x)\mathrm{d}x=\int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x. \end{aligned}$$

综合上两式,(1.4)式得证. □

上述定理还可以推广到两个或两个以上随机变量的函数的情况.

例如,设 Z 是随机变量 X,Y 的函数 $Z=g(X,Y)$ (g 是连续函数),那么,Z 是一个一维随机变量. 若二维随机变量(X,Y)的概率密度为 $f(x,y)$,则有

$$E(Z)=E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y, \tag{1.5}$$

这里设上式右边的积分绝对收敛. 又若(X,Y)为离散型随机变量，其分布律为$P\{X=x_i,Y=y_i\}=p_{ij},i,j=1,2,\cdots$，则有

$$E(Z)=E[g(X,Y)]=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}g(x_i,y_j)p_{ij}, \tag{1.6}$$

这里设上式右边的级数绝对收敛.

例 8　设风速V在$(0,a)$上服从均匀分布，即具有概率密度

$$f(v)=\begin{cases}\dfrac{1}{a}, & 0<v<a,\\ 0, & 其他.\end{cases}$$

又设飞机机翼受到的正压力W是V的函数：$W=kV^2$ ($k>0$，常数)，求W的数学期望.

解　由(1.4)式有

$$E(W)=\int_{-\infty}^{\infty}kv^2f(v)\mathrm{d}v=\int_0^a kv^2\frac{1}{a}\mathrm{d}v=\frac{1}{3}ka^2. \qquad \square$$

例 9　设随机变量(X,Y)的概率密度

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

求数学期望$E(Y)$，$E\left(\dfrac{1}{XY}\right)$.

解　由(1.5)式得

$$\begin{aligned}E(Y)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\mathrm{d}y\mathrm{d}x=\int_1^{\infty}\int_{\frac{1}{x}}^{x}\frac{3}{2x^3y}\mathrm{d}y\mathrm{d}x\\&=\frac{3}{2}\int_1^{\infty}\frac{1}{x^3}\Big[\ln y\Big]_{\frac{1}{x}}^{x}\mathrm{d}x=3\int_1^{\infty}\frac{\ln x}{x^3}\mathrm{d}x\\&=\left[-\frac{3}{2}\frac{\ln x}{x^2}\right]_1^{\infty}+\frac{3}{2}\int_1^{\infty}\frac{1}{x^3}\mathrm{d}x=\frac{3}{4}.\end{aligned}$$

$$E\left(\frac{1}{XY}\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{xy}f(x,y)\mathrm{d}y\mathrm{d}x=\int_1^{\infty}\mathrm{d}x\int_{\frac{1}{x}}^{x}\frac{3}{2x^4y^3}\mathrm{d}y=\frac{3}{5}. \qquad \square$$

例 10　某公司计划开发一种新产品市场，并试图确定该产品的产量. 他们估计出售一件产品可获利m元，而积压一件产品导致n元的损失. 再者，他们预测销售量Y(件)服从指数分布，其概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-y/\theta}, & y>0,\\ 0, & y\leqslant 0.\end{cases}\quad \theta>0,$$

问若要获得利润的数学期望最大，应生产多少件产品(m,n,θ均为已知)？

解　设生产x件，则获利Q是x的函数

$$Q=Q(x)=\begin{cases} mY-n(x-Y), & Y<x, \\ mx, & Y\geqslant x. \end{cases}$$

Q 是随机变量，它是 Y 的函数，其数学期望为

$$\begin{aligned} E(Q) &= \int_0^{\infty} Q f_Y(y)\mathrm{d}y \\ &= \int_0^{x}[my-n(x-y)]\frac{1}{\theta}\mathrm{e}^{-y/\theta}\mathrm{d}y \\ &\quad +\int_x^{\infty} mx\,\frac{1}{\theta}\mathrm{e}^{-y/\theta}\mathrm{d}y \\ &= (m+n)\theta-(m+n)\theta\mathrm{e}^{-x/\theta}-nx. \end{aligned}$$

令
$$\frac{\mathrm{d}}{\mathrm{d}x}E(Q)=(m+n)\mathrm{e}^{-x/\theta}-n=0,$$

得
$$x=-\theta\ln\frac{n}{m+n}.$$

而
$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}E(Q)=\frac{-(m+n)}{\theta}\mathrm{e}^{-x/\theta}<0,$$

故知当 $x=-\theta\ln\dfrac{n}{m+n}$ 时 $E(Q)$ 取极大值，且可知这也是最大值.

例如，若 $f_Y(y)=\begin{cases}\dfrac{1}{10\ 000}\mathrm{e}^{-\frac{y}{10\ 000}}, & y>0, \\ 0, & y\leqslant 0,\end{cases}$

且有 $m=500$ 元，$n=2000$ 元，则

$$x=-10\ 000\ln\frac{2\ 000}{500+2\ 000}=2\ 231.4.$$

取 $x=2231$ 件. □

例 11 某甲与其他三人参与一个项目的竞拍，价格以千美元计，价格高者获胜. 若甲中标，他就将此项目以 10 千美元转让给他人. 可认为其他三人的竞拍价是相互独立的，且都在 7～11 千美元之间均匀分布. 问甲应如何报价才能使获益的数学期望为最大(若甲中标必须将此项目以他自己的报价买下).

解 设 X_1,X_2,X_3 是其他三人的报价，按题意 X_1,X_2,X_3 相互独立，且在区间(7,11)上服从均匀分布. 其分布函数为

$$F(u)=\begin{cases} 0, & u<7, \\ \dfrac{u-7}{4}, & 7\leqslant u<11, \\ 1, & u\geqslant 11. \end{cases}$$

以 Y 记三人最大出价，即 $Y=\max\{X_1,X_2,X_3\}$. Y 的分布函数为

$$F_Y(u)=\begin{cases}0, & u<7,\\ \left(\dfrac{u-7}{4}\right)^3, & 7\leqslant u<11,\\ 1, & u\geqslant 11.\end{cases}$$

若甲的报价为 x，按题意 $7\leqslant x\leqslant 10$，知甲能赢得这一项目的概率为

$$p=P\{Y\leqslant x\}=F_Y(x)=\left(\frac{x-7}{4}\right)^3 \quad (7\leqslant x\leqslant 10).$$

以 $G(x)$ 记甲的赚钱数，$G(x)$ 是一个随机变量，它的分布律为

$G(x)$	$10-x$	0
概率	$\left(\frac{x-7}{4}\right)^3$	$1-\left(\frac{x-7}{4}\right)^3$

于是甲赚钱数的数学期望为

$$E[G(x)]=\left(\frac{x-7}{4}\right)^3(10-x).$$

令 $$\frac{\mathrm{d}}{\mathrm{d}x}E[G(x)]=\frac{1}{4^3}[(x-7)^2(37-4x)]=0,$$

得 $$x=37/4, x=7\ (舍去).$$

又知 $$\left.\frac{\mathrm{d}^2}{\mathrm{d}x^2}E[G(x)]\right|_{x=37/4}<0.$$

故知当甲的报价为 $x=37/4$ 千美元时，他赚钱数的数学期望达到极大值，还可知这也是最大值. □

现在来证明数学期望的几个重要性质①（以下设所遇到的随机变量的数学期望存在）.

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

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

这一性质可以推广到任意有限个随机变量之和的情况.

4° 设 X,Y 是相互独立的随机变量，则有

$$E(XY)=E(X)E(Y).$$

这一性质可以推广到任意有限个相互独立的随机变量之积的情况.

证 1°、2°由读者自己证明. 我们来证 3°和 4°.

---

① 这里我们只对连续型随机变量的情况加以证明，读者只要将证明中的“积分”用“和式”代替，就能得到离散型随机变量情况的证明.

设二维随机变量(X,Y)的概率密度为$f(x,y)$. 其边缘概率密度为$f_X(x)$，$f_Y(y)$. 由(1.5)式

$$\begin{aligned}E(X+Y)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x\mathrm{d}y+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\mathrm{d}x\mathrm{d}y\\&=E(X)+E(Y).\end{aligned}$$

3°得证.

又若X和Y相互独立，

$$\begin{aligned}E(XY)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y\\&=\left[\int_{-\infty}^{\infty}xf_X(x)\mathrm{d}x\right]\left[\int_{-\infty}^{\infty}yf_Y(y)\mathrm{d}y\right]=E(X)E(Y).\end{aligned}$$

4°得证. □

例 12 一民航送客车载有 20 位旅客自机场开出，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，并设各位旅客是否下车相互独立).

解 引入随机变量

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

易知
$$X=X_1+X_2+\cdots+X_{10}.$$
现在来求$E(X)$.

按题意，任一旅客在第i站不下车的概率为$\frac{9}{10}$，因此 20 位旅客都不在第i站下车的概率为$\left(\frac{9}{10}\right)^{20}$，在第$i$站有人下车的概率为$1-\left(\frac{9}{10}\right)^{20}$，也就是

$$P\{X_i=0\}=\left(\frac{9}{10}\right)^{20},P\{X_i=1\}=1-\left(\frac{9}{10}\right)^{20},i=1,2,\cdots,10.$$

由此

$$E(X_i)=1-\left(\frac{9}{10}\right)^{20},i=1,2,\cdots,10.$$

进而

$$\begin{aligned}E(X)&=E(X_1+X_2+\cdots+X_{10})\\&=E(X_1)+E(X_2)+\cdots+E(X_{10})\\&=10\left[1-\left(\frac{9}{10}\right)^{20}\right]=8.784(\text{次}).\end{aligned}$$

本题是将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来求数学期望的，这种处理方法具有一定的普遍意义. □

例 13 设一电路中电流 I (A)与电阻 R (Ω)是两个相互独立的随机变量，其概率密度分别为

$$g(i)=\begin{cases}2i, & 0\leqslant i\leqslant 1,\\ 0, & \text{其他},\end{cases} \quad h(r)=\begin{cases}\dfrac{r^2}{9}, & 0\leqslant r\leqslant 3,\\ 0, & \text{其他}.\end{cases}$$

试求电压 $V=IR$ 的均值.

解

$$\begin{aligned}E(V)&=E(IR)=E(I)E(R)\\&=\left[\int_{-\infty}^{\infty} ig(i)\mathrm{d}i\right]\left[\int_{-\infty}^{\infty} rh(r)\mathrm{d}r\right]\\&=\left(\int_0^1 2i^2\mathrm{d}i\right)\left(\int_0^3 \frac{r^3}{9}\mathrm{d}r\right)=\frac{3}{2}(\mathrm{V}).\end{aligned}$$ □

§2 方　　差

先从例子说起. 例如，有一批灯泡，知其平均寿命是 $E(X)=1\ 000$(小时). 仅由这一指标我们还不能判定这批灯泡的质量好坏. 事实上，有可能其中绝大部分灯泡的寿命都在 950～1 050 小时；也有可能其中约有一半是高质量的，它们的寿命大约有 1 300 小时，另一半却是质量很差的，其寿命大约只有 700 小时. 为要评定这批灯泡质量的好坏，还需进一步考察灯泡寿命 X 与其均值 $E(X)=1\ 000$ 的偏离程度. 若偏离程度较小，表示质量比较稳定. 从这个意义上来说，我们认为质量较好. 前面也曾提到在检验棉花的质量时，既要注意纤维的平均长度，还要注意纤维长度与平均长度的偏离程度. 由此可见，研究随机变量与其均值的偏离程度是十分必要的. 那么，用怎样的量去度量这个偏离程度呢？容易看到

$$E\{|X-E(X)|\}$$

能度量随机变量与其均值 $E(X)$ 的偏离程度. 但由于上式带有绝对值，运算不方便，为运算方便起见，通常用量

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

来度量随机变量 X 与其均值 $E(X)$ 的偏离程度.

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

$$D(X)=\mathrm{Var}(X)=E\{[X-E(X)]^2\}. \tag{2.1}$$

在应用上还引入量 $\sqrt{D(X)}$，记为 $\sigma(X)$，称为**标准差**或**均方差**.

按定义，随机变量 X 的方差表达了 X 的取值与其数学期望的偏离程度．若 $D(X)$ 较小意味着 X 的取值比较集中在 $E(X)$ 的附近，反之，若 $D(X)$ 较大则表示 X 的取值较分散．因此，$D(X)$ 是刻画 X 取值分散程度的一个量，它是衡量 X 取值分散程度的一个尺度．

由定义知，方差实际上就是随机变量 X 的函数 $g(X)=(X-E(X))^2$ 的数学期望．于是对于离散型随机变量，按(1.3)式有

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2p_k, \tag{2.2}$$

其中 $P\{X=x_k\}=p_k, k=1,2,\cdots$ 是 X 的分布律．

对于连续型随机变量，按(1.4)式有

$$D(X)=\int_{-\infty}^{\infty}[x-E(X)]^2f(x)\mathrm{d}x, \tag{2.3}$$

其中 $f(x)$ 是 X 的概率密度．

随机变量 X 的方差可按下列公式计算．

$$D(X)=E(X^2)-[E(X)]^2. \tag{2.4}$$

证　由数学期望的性质 1°，2°，3°得

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

□

例 1　设随机变量 X 具有数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2\neq 0$．记

$$X^*=\frac{X-\mu}{\sigma},$$

则

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

$$\begin{aligned}D(X^*)&=E(X^{*2})-[E(X^*)]^2=E\left[\left(\frac{X-\mu}{\sigma}\right)^2\right]\\&=\frac{1}{\sigma^2}E[(X-\mu)^2]=\frac{\sigma^2}{\sigma^2}=1.\end{aligned}$$

即 $X^*=\dfrac{X-\mu}{\sigma}$ 的数学期望为 0，方差为 1．X^* 称为 X 的**标准化变量**．□

例 2　设随机变量 X 具有(0－1)分布，其分布律为

$$P\{X=0\}=1-p,\quad P\{X=1\}=p.$$

求 $D(X)$．

解

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

由(2.4)式

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

□

例 3　设随机变量 $X\sim\pi(\lambda)$，求 $D(X)$.

解　随机变量 X 的分布律为

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,2,\cdots,\quad \lambda>0.$$

上节例 6 已算得 $E(X)=\lambda$，而

$$\begin{aligned}E(X^2)&=E[X(X-1)+X]=E[X(X-1)]+E(X)\\&=\sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k e^{-\lambda}}{k!}+\lambda=\lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda\\&=\lambda^2 e^{-\lambda}e^{\lambda}+\lambda=\lambda^2+\lambda,\end{aligned}$$

所以方差

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

由此可知，泊松分布的数学期望与方差相等，都等于参数 λ. 因为泊松分布只含一个参数 λ，只要知道它的数学期望或方差就能完全确定它的分布了。 □

例 4　设随机变量 $X\sim U(a,b)$，求 $D(X)$.

解　X 的概率密度为

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

上节例 7 已算得 $E(X)=\dfrac{a+b}{2}$. 方差为

$$\begin{aligned}D(X)&=E(X^2)-[E(X)]^2\\&=\int_a^b x^2\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}.\end{aligned}$$ □

例 5　设随机变量 X 服从指数分布，其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

其中 $\theta>0$，求 $E(X)$，$D(X)$.

解

$$\begin{aligned}E(X)&=\int_{-\infty}^{\infty}xf(x)dx=\int_0^{\infty}x\frac{1}{\theta}e^{-x/\theta}dx\\&=-xe^{-x/\theta}\Big|_0^{\infty}+\int_0^{\infty}e^{-x/\theta}dx=\theta,\\E(X^2)&=\int_{-\infty}^{\infty}x^2f(x)dx=\int_0^{\infty}x^2\frac{1}{\theta}e^{-x/\theta}dx\\&=-x^2e^{-x/\theta}\Big|_0^{\infty}+\int_0^{\infty}2xe^{-x/\theta}dx=2\theta^2,\end{aligned}$$

于是

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

即有
$$E(X)=\theta,\quad D(X)=\theta^2.$$ □

现在来证明方差的几个重要性质(以下设所遇到的随机变量其方差存在).

1° 设 C 是常数,则 $D(C)=0$.

2° 设 X 是随机变量,C 是常数,则有
$$D(CX)=C^2D(X),\qquad D(X+C)=D(X).$$

3° 设 X,Y 是两个随机变量,则有
$$D(X+Y)=D(X)+D(Y)+2E\{(X-E(X))(Y-E(Y))\}.\tag{2.5}$$
特别,若 X,Y 相互独立,则有
$$D(X+Y)=D(X)+D(Y).\tag{2.6}$$
这一性质可以推广到任意有限多个相互独立的随机变量之和的情况.

4° $D(X)=0$ 的充要条件是 X 以概率 1 取常数 $E(X)$,即
$$P\{X=E(X)\}=1.$$

证 1° $D(C)=E\{[C-E(C)]^2\}=0$.

2° $D(CX)=E\{[CX-E(CX)]^2\}=C^2E\{[X-E(X)]^2\}=C^2D(X)$.

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

3°
$$\begin{aligned}D(X+Y)&=E\{[(X+Y)-E(X+Y)]^2\}\\&=E\{[(X-E(X))+(Y-E(Y))]^2\}\\&=E\{(X-E(X))^2\}+E\{(Y-E(Y))^2\}\\&\quad+2E\{[(X-E(X)][Y-E(Y)]\}\\&=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}.\end{aligned}$$
上式右端第三项:
$$\begin{aligned}&2E\{[(X-E(X)][Y-E(Y)]\}\\&=2E\{XY-XE(Y)-YE(X)+E(X)E(Y)\}\\&=2\{E(XY)-E(X)E(Y)-E(Y)E(X)+E(X)E(Y)\}\\&=2\{E(XY)-E(X)E(Y)\}.\end{aligned}$$
若 X,Y 相互独立,由数学期望的性质 4°知道上式右端为 0,于是
$$D(X+Y)=D(X)+D(Y).$$

4° 充分性. 设 $P\{X=E(X)\}=1$,则有 $P\{X^2=[E(X)]^2\}=1$,于是
$$D(X)=E(X^2)-[E(X)]^2=0.$$
必要性的证明写在切比雪夫不等式证明的后面. □

例 6 设随机变量 $X\sim b(n,p)$,求 $E(X),D(X)$.

解 由二项分布的定义知,随机变量 X 是 n 重伯努利试验中事件 A 发生的次数,且在每次试验中 A 发生的概率为 p. 引入随机变量
$$X_k=\begin{cases}1, & A\text{ 在第 }k\text{ 次试验发生},\\0, & A\text{ 在第 }k\text{ 次试验不发生},\end{cases}\quad k=1,2,\cdots,n.$$

易知
$$X=X_1+X_2+\cdots+X_n, \tag{2.7}$$
由于 X_k 只依赖于第 k 次试验，而各次试验相互独立，于是 $X_1,X_2,\cdots,X_n$ 相互独立，又知 $X_k,k=1,2,\cdots,n$ 服从同一(0－1)分布

X_k	0	1
p_k	$1-p$	p

(2.7)式表明以 n,p 为参数的二项分布变量，可分解成为 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．

由例 2 知 $E(X_k)=p,D(X_k)=p(1-p),k=1,2,\cdots,n$. 故知
$$E(X)=E\Big(\sum_{k=1}^{n}X_k\Big)=\sum_{k=1}^{n}E(X_k)=np.$$
又由于 $X_1,X_2,\cdots,X_n$ 相互独立，得
$$D(X)=D\Big(\sum_{k=1}^{n}X_k\Big)=\sum_{k=1}^{n}D(X_k)=np(1-p).$$
即
$$E(X)=np,\quad D(X)=np(1-p).$$
□

例 7　设随机变量 $X\sim N(\mu,\sigma^2)$，求 $E(X),D(X)$.

解　先求标准正态变量
$$Z=\frac{X-\mu}{\sigma}$$
的数学期望和方差. Z 的概率密度为
$$\varphi(t)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2},$$
于是
$$E(Z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t\mathrm{e}^{-t^2/2}\mathrm{d}t=\frac{-1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\Big|_{-\infty}^{\infty}=0,$$
$$D(Z)=E(Z^2)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2\mathrm{e}^{-t^2/2}\mathrm{d}t$$
$$=\frac{-1}{\sqrt{2\pi}}t\mathrm{e}^{-t^2/2}\Big|_{-\infty}^{\infty}+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-t^2/2}\mathrm{d}t=1.$$
因　$X=\mu+\sigma Z$，即得
$$E(X)=E(\mu+\sigma Z)=\mu,$$
$$D(X)=D(\mu+\sigma Z)=D(\sigma Z)=\sigma^2D(Z)=\sigma^2.$$
这就是说，正态分布的概率密度中的两个参数 μ 和 σ 分别就是该分布的数学期望和均方差，因而正态分布完全可由它的数学期望和方差所确定.

再者，由上一章§5 中例 1 知道，若 $X_i\sim N(\mu_i,\sigma_i^2),i=1,2,\cdots,n$，且它们相互独立，则它们的线性组合：$C_1X_1+C_2X_2+\cdots+C_nX_n$ （$C_1,C_2,\cdots,C_n$ 是不全为 0 的常数）仍然服从正态分布，于是由数学期望和方差的性质知道

$$C_1X_1+C_2X_2+\cdots+C_nX_n\sim N\Big(\sum_{i=1}^{n}C_i\mu_i,\sum_{i=1}^{n}C_i^2\sigma_i^2\Big) \tag{2.8}$$

这一重要结果.

例如,若 $X\sim N(1,3)$,$Y\sim N(2,4)$且 X,Y 相互独立,则 $Z=2X-3Y$ 也服从正态分布,而 $E(Z)=2\times1-3\times2=-4$,$D(Z)=D(2X-3Y)=4D(X)+9D(Y)=48$.故有 $Z\sim N(-4,48)$. □

例 8　设活塞的直径(以 cm 计)$X\sim N(22.40,0.03^2)$,气缸的直径 $Y\sim N(22.50,0.04^2)$,X,Y 相互独立.任取一只活塞,任取一只气缸,求活塞能装入气缸的概率.

解　按题意需求 $P\{X<Y\}=P\{X-Y<0\}$.由于

$$X-Y\sim N(-0.10,0.002\,5),$$

故有

$$\begin{aligned}P\{X<Y\}&=P\{X-Y<0\}\\&=P\left\{\frac{(X-Y)-(-0.10)}{\sqrt{0.002\,5}}<\frac{0-(-0.10)}{\sqrt{0.002\,5}}\right\}\\&=\Phi\left(\frac{0.10}{0.05}\right)=\Phi(2)=0.977\,2.\end{aligned}$$

□

下面介绍一个重要的不等式.

定理　设随机变量 X 具有数学期望 $E(X)=\mu$,方差 $D(X)=\sigma^2$,则对于任意正数 ε,不等式

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\frac{\sigma^2}{\varepsilon^2} \tag{2.9}$$

成立.

这一不等式称为**切比雪夫(Chebyshev)不等式**.

证　我们只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(如图 4—2)

$$\begin{aligned}P\{|X-\mu|\geqslant\varepsilon\}&=\int_{|x-\mu|\geqslant\varepsilon}f(x)\mathrm{d}x\leqslant\int_{|x-\mu|\geqslant\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)\mathrm{d}x\\&\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}(x-\mu)^2f(x)\mathrm{d}x=\frac{\sigma^2}{\varepsilon^2}.\end{aligned}$$

□

切比雪夫不等式也可以写成如下的形式:

$$P\{|X-\mu|<\varepsilon\}\geqslant1-\frac{\sigma^2}{\varepsilon^2}. \tag{2.10}$$

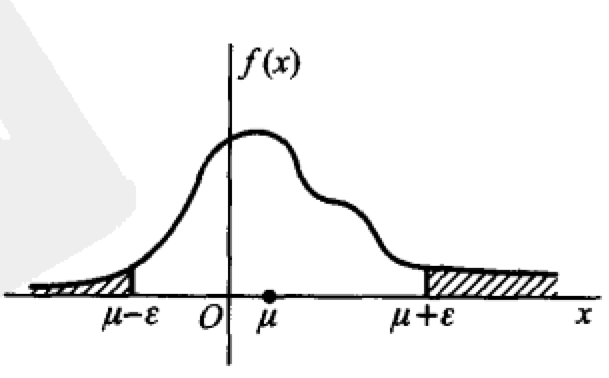


图 4—2

切比雪夫不等式给出了在随机变量的分布未知,而只知道 $E(X)$ 和 $D(X)$ 的情况下估计概率 $P\{|X-E(X)|<\varepsilon\}$ 的界限.例

如在(2.9)式中分别取 $\varepsilon=3\sqrt{D(X)}$，$4\sqrt{D(X)}$ 得到

$$P\{|X-E(X)|<3\sqrt{D(X)}\}\geqslant 0.8889,$$

$$P\{|X-E(X)|<4\sqrt{D(X)}\geqslant 0.9375.$$

这个估计是比较粗糙的①，如果已经知道随机变量的分布时，那么所需求的概率可以确切地计算出来，也就没有必要利用这一不等式来作估计了.

方差性质 4°必要性的证明：

设 $D(X)=0$，要证 $P\{X=E(X)\}=1$.

证　用反证法　假设 $P\{X=E(X)\}<1$，则对于某一个数 $\varepsilon>0$，有 $P\{|X-E(X)|\geqslant\varepsilon\}>0$，但由切比雪夫不等式，对于任意 $\varepsilon>0$，由(2.9)式因 $\sigma^2=0$，有

$$P\{|X-E(X)|\geqslant\varepsilon\}=0,$$

矛盾，于是 $P\{X=E(X)\}=1$. □

在书末附表 1 中列出了多种常用的随机变量的数学期望和方差，供读者查用.

§3　协方差及相关系数

对于二维随机变量 (X,Y)，我们除了讨论 X 与 Y 的数学期望和方差以外，还需讨论描述 X 与 Y 之间相互关系的数字特征. 本节讨论有关这方面的数字特征.

在本章§2 方差性质 3°的证明中，我们已经看到，如果两个随机变量 X 和 Y 是相互独立的，则

$$E\{[X-E(X)][Y-E(Y)]\}=0.$$

这意味着当 $E\{[X-E(X)][Y-E(Y)]\}\neq 0$ 时，X 与 Y 不相互独立，而是存在着一定的关系的.

定义　量 $E\{[X-E(X)][Y-E(Y)]\}$ 称为随机变量 X 与 Y 的**协方差**. 记为 $\mathrm{Cov}(X,Y)$，即

$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}.$$

而

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

称为随机变量 X 与 Y 的**相关系数**.

由定义，即知

---

① 例如若 $X\sim U(0,8)$，则有 $E(X)=4$，$D(X)=16/3$，由切比雪夫不等式(2.10)，$P\{|X-4|<4\}\geqslant 1-1/3=2/3$，但准确的结果是 $P\{|X-4|<4\}=1$.

$$\operatorname{Cov}(X,Y)=\operatorname{Cov}(Y,X),\quad \operatorname{Cov}(X,X)=D(X).$$

由上述定义及(2.5)式知道，对于任意两个随机变量 X 和 Y，下列等式成立：

$$D(X+Y)=D(X)+D(Y)+2\operatorname{Cov}(X,Y). \tag{3.1}$$

将 $\operatorname{Cov}(X,Y)$ 的定义式展开，易得

$$\operatorname{Cov}(X,Y)=E(XY)-E(X)E(Y). \tag{3.2}$$

我们常常利用这一式子计算协方差.

协方差具有下述性质：

1° $\operatorname{Cov}(aX,bY)=ab\operatorname{Cov}(X,Y)$，$a,b$ 是常数.

2° $\operatorname{Cov}(X_1+X_2,Y)=\operatorname{Cov}(X_1,Y)+\operatorname{Cov}(X_2,Y)$.

(证明由读者自己来完成.)

下面我们来推导 ρ_{XY} 的两条重要性质，并说明 ρ_{XY} 的含义.

考虑以 X 的线性函数 $a+bX$ 来近似表示 Y. 我们以均方误差

$$\begin{aligned} e &= E[(Y-(a+bX))^2] \\ &= E(Y^2)+b^2E(X^2)+a^2-2bE(XY)+2abE(X)-2aE(Y) \end{aligned} \tag{3.3}$$

来衡量以 $a+bX$ 近似表达 Y 的好坏程度. e 的值越小表示 $a+bX$ 与 Y 的近似程度越好. 这样，我们就取 a,b 使 e 取到最小. 下面就来求最佳近似式 $a+bX$ 中的 a,b. 为此，将 e 分别关于 a,b 求偏导数，并令它们等于零，得

$$\begin{cases} \dfrac{\partial e}{\partial a}=2a+2bE(X)-2E(Y)=0, \\ \dfrac{\partial e}{\partial b}=2bE(X^2)-2E(XY)+2aE(X)=0. \end{cases}$$

解得
$$b_0=\frac{\operatorname{Cov}(X,Y)}{D(X)},$$

$$a_0=E(Y)-b_0E(X)=E(Y)-E(X)\frac{\operatorname{Cov}(X,Y)}{D(X)}.$$

将 a_0,b_0 代入(3.3)式得

$$\begin{aligned} \min_{a,b} E\{[Y-(a+bX)]^2\} &= E\{[Y-(a_0+b_0X)]^2\} \\ &= (1-\rho_{XY}^2)D(Y)\text{①}. \end{aligned} \tag{3.4}$$

---

① $$\begin{aligned} E[(Y-(a_0+b_0X))^2] &= D[Y-a_0-b_0X]+[E(Y-a_0-b_0X)]^2 \\ &= D(Y-b_0X)-\left[\frac{1}{2}\frac{\partial e}{\partial a}\bigg|_{\substack{a=a_0\\ b=b_0}}\right]^2=D(Y-b_0X)+0 \\ &= D(Y)+b_0^2D(X)-2b_0\operatorname{Cov}(X,Y)=D(Y)+\frac{\operatorname{Cov}^2(X,Y)}{D(X)}-2\frac{\operatorname{Cov}^2(X,Y)}{D(X)} \\ &= D(Y)\left[1-\frac{\operatorname{Cov}^2(X,Y)}{D(X)D(Y)}\right]=(1-\rho_{XY}^2)D(Y). \end{aligned}$$

由(3.4)式容易得到下述定理：

定理　1° $|\rho_{XY}|\leqslant 1$.

2° $|\rho_{XY}|=1$ 的充要条件是，存在常数 a,b 使

$$P\{Y=a+bX\}=1.$$

证　1° 由(3.4)式与 $E\{[Y-(a_0+b_0X)]^2\}$ 及 $D(Y)$ 的非负性，得知 $1-\rho_{XY}^2\geqslant 0$，亦即 $|\rho_{XY}|\leqslant 1$.

2° 若 $|\rho_{XY}|=1$ 由(3.4)式得

$$E\{[Y-(a_0+b_0X)]^2\}=0.$$

从而　$0=E\{[Y-(a_0+b_0X)]^2\}=D[Y-(a_0+b_0X)]+[E(Y-(a_0+b_0X))]^2$，

故有
$$D[Y-(a_0+b_0X)]=0,$$
$$E[Y-(a_0+b_0X)]=0.$$

又由方差的性质 4°知

$$P\{Y-(a_0+b_0X)=0\}=1，即\ P\{Y=a_0+b_0X\}=1.$$

反之，若存在常数 a^*,b^* 使

$$P\{Y=a^*+b^*X\}=1，即\ P\{Y-(a^*+b^*X)=0\}=1,$$

于是
$$P\{[Y-(a^*+b^*X)]^2=0\}=1.$$

即得
$$E\{[Y-(a^*+b^*X)]^2\}=0.$$

故有
$$0=E\{[Y-(a^*+b^*X)]^2\}\geqslant\min_{a,b}E\{[Y-(a+bX)]^2\}$$
$$=E\{[Y-(a_0+b_0X)]^2\}=(1-\rho_{XY}^2)D(Y).$$

即得
$$|\rho_{XY}|=1.$$
□

由(3.4)知，均方误差 e 是 $|\rho_{XY}|$ 的严格单调减少函数，这样 ρ_{XY} 的含义就很明显了. 当 $|\rho_{XY}|$ 较大时 e 较小，表明 X,Y(就线性关系来说)联系较紧密. 特别当 $|\rho_{XY}|=1$ 时，由定理中的 2°，X,Y 之间以概率 1 存在着线性关系. 于是 ρ_{XY} 是一个可以用来表征 X,Y 之间线性关系紧密程度的量. 当 $|\rho_{XY}|$ 较大时，我们通常说 X,Y 线性相关的程度较好；当 $|\rho_{XY}|$ 较小时，我们说，X,Y 线性相关的程度较差.

当 $\rho_{XY}=0$ 时，称 X 和 Y **不相关**.

假设随机变量 X,Y 的相关系数 ρ_{XY} 存在. 当 X 和 Y 相互独立时，由数学期望的性质 4°及(3.2)式知 $\mathrm{Cov}(X,Y)=0$，从而 $\rho_{XY}=0$，即 X,Y 不相关. 反之，若 X,Y 不相关，X 和 Y 却不一定相互独立(见例 1). 上述情况，从"不相关"和"相互独立"的含义来看是明显的. 这是因为不相关只是就线性关系来说的，而相互独立是就一般关系而言的.

不过，从例 2 可以看到，当(X,Y)服从二维正态分布时，X 和 Y 不相关与 X 和 Y 相互独立是等价的.

例 1　设(X,Y)的分布律为

X \ Y	-2	-1	1	2	$P\{Y=i\}$
1	0	1/4	1/4	0	1/2
4	1/4	0	0	1/4	1/2
$P\{X=i\}$	1/4	1/4	1/4	1/4	1

易知 $E(X)=0,E(Y)=5/2,E(XY)=0$，于是 $\rho_{XY}=0$，X,Y 不相关. 这表示 X,Y 不存在线性关系. 但，$P\{X=-2,Y=1\}=0\neq P\{X=-2\}P\{Y=1\}$，知 X,Y 不是相互独立的. 事实上，X 和 Y 具有关系：$Y=X^2$，Y 的值完全可由 X 的值所确定. □

例 2 设 (X,Y) 服从二维正态分布，它的概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

我们来求 X 和 Y 的相关系数.

在第三章§2例3中已经知道(X,Y)的边缘概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},-\infty<x<\infty,$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}},-\infty<y<\infty.$$

故知 $E(X)=\mu_1,E(Y)=\mu_2,D(X)=\sigma_1^2,D(Y)=\sigma_2^2$. 而

$$\begin{aligned}\operatorname{Cov}(X,Y)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)\,\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)\\&\quad\times\exp\left[\frac{-1}{2(1-\rho^2)}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]\mathrm{d}y\mathrm{d}x.\end{aligned}$$

令 $t=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right),u=\frac{x-\mu_1}{\sigma_1}$，则有

$$\begin{aligned}\operatorname{Cov}(X,Y)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sigma_1\sigma_2\sqrt{1-\rho^2}\,tu+\rho\sigma_1\sigma_2u^2)e^{-(u^2+t^2)/2}\mathrm{d}t\mathrm{d}u\\&=\frac{\rho\sigma_1\sigma_2}{2\pi}\left(\int_{-\infty}^{\infty}u^2e^{-\frac{u^2}{2}}\mathrm{d}u\right)\left(\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}\mathrm{d}t\right)\\&\quad+\frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\left(\int_{-\infty}^{\infty}ue^{-\frac{u^2}{2}}\mathrm{d}u\right)\left(\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}\mathrm{d}t\right)\end{aligned}$$

$$=\frac{\rho\sigma_1\sigma_2}{2\pi}\sqrt{2\pi}\cdot\sqrt{2\pi},$$

即有
$$\mathrm{Cov}(X,Y)=\rho\sigma_1\sigma_2.$$

于是
$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\rho.$$
□

这就是说，二维正态随机变量(X,Y)的概率密度中的参数ρ就是X和Y的相关系数，因而二维正态随机变量的分布完全可由X,Y各自的数学期望、方差以及它们的相关系数所确定.

在第三章§4中已经讲过，若(X,Y)服从二维正态分布，那么X和Y相互独立的充要条件为$\rho=0$. 现在知道$\rho=\rho_{XY}$，故知对于二维正态随机变量(X,Y)来说，X和Y不相关与X和Y相互独立是等价的.

§4　矩、协方差矩阵

本节先介绍随机变量的另外几个数字特征. 设(X,Y)是二维随机变量.

定义　设X和Y是随机变量，若
$$E(X^k),\quad k=1,2,\cdots$$
存在，称它为X的k**阶原点矩**，简称k**阶矩**.

若
$$E\{[X-E(X)]^k\},k=2,3,\cdots$$
存在，称它为X的k**阶中心矩**.

若
$$E(X^kY^l),\quad k,l=1,2,\cdots$$
存在，称它为X和Y的$k+l$**阶混合矩**.

若
$$E\{[X-E(X)]^k[Y-E(Y)]^l\},\quad k,l=1,2,\cdots$$
存在，称它为X和Y的$k+l$**阶混合中心矩**.

显然，X的数学期望$E(X)$是X的一阶原点矩，方差$D(X)$是X的二阶中心矩，协方差$\mathrm{Cov}(X,Y)$是X和Y的二阶混合中心矩.

下面介绍n维随机变量的协方差矩阵. 先从二维随机变量讲起.

二维随机变量(X_1,X_2)有四个二阶中心矩（设它们都存在），分别记为
$$c_{11}=E\{[X_1-E(X_1)]^2\},$$
$$c_{12}=E\{[X_1-E(X_1)][X_2-E(X_2)]\},$$
$$c_{21}=E\{[X_2-E(X_2)][X_1-E(X_1)]\},$$
$$c_{22}=E\{[X_2-E(X_2)]^2\}.$$
将它们排成矩阵的形式
$$\begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix}.$$

这个矩阵称为随机变量(X_1,X_2)的**协方差矩阵**.

设n维随机变量$(X_1,X_2,\cdots,X_n)$的二阶混合中心矩

$$c_{ij}=\mathrm{Cov}(X_i,X_j)=E\{[X_i-E(X_i)][X_j-E(X_j)]\},i,j=1,2,\cdots,n$$

都存在,则称矩阵

$$\boldsymbol{C}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为n维随机变量$(X_1,X_2,\cdots,X_n)$的**协方差矩阵**.由于$c_{ij}=c_{ji}$ $(i\neq j;i,j=1,2,\cdots,n)$,因而上述矩阵是一个对称矩阵.

一般,n维随机变量的分布是不知道的,或者是太复杂,以致在数学上不易处理,因此在实际应用中协方差矩阵就显得重要了.

本节的最后,介绍n维正态随机变量的概率密度.我们先将二维正态随机变量的概率密度改写成另一种形式,以便将它推广到n维随机变量的场合中去.二维正态随机变量(X_1,X_2)的概率密度为

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}.$$

现在将上式中花括号内的式子写成矩阵形式,为此引入下面的列矩阵

$$\boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\quad \boldsymbol{\mu}=\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

(X_1,X_2)的协方差矩阵为

$$\boldsymbol{C}=\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}=\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

它的行列式$\det \boldsymbol{C}=\sigma_1^2\sigma_2^2(1-\rho^2)$,$\boldsymbol{C}$的逆矩阵为

$$\boldsymbol{C}^{-1}=\frac{1}{\det \boldsymbol{C}}\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}.$$

经过计算可知(这里矩阵$(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}$是$(\boldsymbol{X}-\boldsymbol{\mu})$的转置矩阵)

$$\begin{aligned}&(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\\ &=\frac{1}{\det \boldsymbol{C}}(x_1-\mu_1 \quad x_2-\mu_2)\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}\\ &=\frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right].\end{aligned}$$

于是(X_1,X_2)的概率密度可写成

$$f(x_1,x_2)=\frac{1}{(2\pi)^{2/2}(\det \boldsymbol{C})^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right\}.$$

上式容易推广到 n 维正态随机变量$(X_1,X_2,\cdots,X_n)$的情况.

引入列矩阵

$$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\text{和}\quad\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}=\begin{pmatrix}E(X_1)\\E(X_2)\\\vdots\\E(X_n)\end{pmatrix}.$$

n 维正态随机变量$(X_1,X_2,\cdots,X_n)$的概率密度定义为

$$f(x_1,x_2,\cdots,x_n)$$
$$=\frac{1}{(2\pi)^{n/2}(\det \boldsymbol{C})^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right\},$$

其中 $\boldsymbol{C}$ 是$(X_1,X_2,\cdots,X_n)$的协方差矩阵.

n 维正态随机变量具有以下四条重要性质(证略):

1° n 维正态随机变量$(X_1,X_2,\cdots,X_n)$的每一个分量 $X_i,i=1,2,\cdots,n$ 都是正态随机变量;反之,若 $X_1,X_2,\cdots,X_n$ 都是正态随机变量,且相互独立,则$(X_1,X_2,\cdots,X_n)$是 n 维正态随机变量.

2° n 维随机变量$(X_1,X_2,\cdots,X_n)$服从 n 维正态分布的充要条件是 $X_1,X_2,\cdots,X_n$ 的任意的线性组合

$$l_1X_1+l_2X_2+\cdots+l_nX_n$$

服从一维正态分布(其中 $l_1,l_2,\cdots,l_n$ 不全为零).

3° 若$(X_1,X_2,\cdots,X_n)$服从 n 维正态分布,设 $Y_1,Y_2,\cdots,Y_k$ 是 $X_j\ (j=1,2,\cdots,n)$的线性函数,则$(Y_1,Y_2,\cdots,Y_k)$也服从多维正态分布.

这一性质称为正态变量的线性变换不变性.

4° 设$(X_1,X_2,\cdots,X_n)$服从 n 维正态分布,则“$X_1,X_2,\cdots,X_n$ 相互独立”与“$X_1,X_2,\cdots,X_n$ 两两不相关”是等价的.

n 维正态分布在随机过程和数理统计中常会遇到.

小结

随机变量的数字特征是由随机变量的分布确定的,能描述随机变量某一个方面的特征的常数. 最重要的数字特征是数学期望和方差. 数学期望 $E(X)$描述随机变量 X 取值的平均大小,方差 $D(X)=E\{[X-E(X)]^2\}$描述随机变量 X 与它自己的数学期望 $E(X)$的偏离程度. 数学期望和方差在应用和理论上都非常重要.

要掌握随机变量的函数 $Y=g(X)$的数学期望 $E(Y)=E[g(X)]$的计算公式(1.3)和

(1.4). 这两个公式的意义在于当我们求 $E(Y)$ 时,不必先求出 $Y=g(X)$ 的分布律或概率密度,而只需利用 X 的分布律或概率密度就可以了,这样做的好处是明显的.

要掌握数学期望和方差的性质. 提请读者注意的是:

(1) 当 X_1,X_2 独立或 X_1,X_2 不相关时,才有 $E(X_1X_2)=E(X_1)E(X_2)$;

(2) 设 C 为常数,则有 $D(CX)=C^2D(X)$,右边的系数是 C^2,不是 C;

(3) $D(X_1+X_2)=D(X_1)+D(X_2)+2\mathrm{Cov}(X_1,X_2)$,当 X_1,X_2 独立或 X_1,X_2 不相关时才有

$$D(X_1+X_2)=D(X_1)+D(X_2).$$

例如,若 X_1,X_2 独立,则有 $D(2X_1-3X_2)=4D(X_1)+9D(X_2)$.

相关系数 ρ_{XY} 有时也称为线性相关系数,它是一个可以用来描述随机变量 (X,Y) 的两个分量 X,Y 之间的线性关系紧密程度的数字特征. 当 $|\rho_{XY}|$ 较小时 X,Y 的线性相关的程度较差;当 $\rho_{XY}=0$ 时称 X,Y 不相关. 不相关是指 X,Y 之间不存在线性关系,X,Y 不相关,它们还可能存在除线性关系之外的关系(参见§3 例 1). 又由于 X,Y 相互独立是指 X,Y 的一般关系而言的,因此有以下的结论:X,Y 相互独立则 X,Y 一定不相关;反之,若 X,Y 不相关则 X,Y 不一定相互独立.

特别,对于二维正态随机变量 (X,Y),X 和 Y 不相关与 X 和 Y 相互独立是等价的. 而二元正态随机变量的相关系数 ρ_{XY} 就是参数 ρ. 于是,用"$\rho=0$"是否成立来检验 X,Y 是否相互独立是很方便的.

切比雪夫不等式给出了在随机变量 X 的分布未知,只知道 $E(X)$ 和 $D(X)$ 的情况下,对事件 $\{|X-E(X)|\leqslant\varepsilon\}$ 概率的下限的估计.

■重要术语及主题

数学期望　随机变量函数的数学期望　数学期望的性质

方差　标准差　方差的性质　标准化的随机变量　协方差　相关系数　相关系数的性质

X,Y 不相关　切比雪夫不等式　几种重要分布的数学期望和方差　矩　协方差矩阵

习题

1. (1) 在下列句子中随机地取一个单词,以 X 表示取到的单词所包含的字母个数,写出 X 的分布律并求 $E(X)$.

"THE GIRL PUT ON HER BEAUTIFUL RED HAT".

(2) 在上述句子的 30 个字母中随机地取一个字母,以 Y 表示取到的字母所在单词所包含的字母数,写出 Y 的分布律并求 $E(Y)$.

(3) 一人掷骰子,如得 6 点则掷第 2 次,此时得分为 6+第二次得到的点数;否则得分为他第一次掷得的点数,且不能再掷,求得分 X 的分布律及 $E(X)$.

2. 某产品的次品率为 0.1,检验员每天检验 4 次. 每次随机地取 10 件产品进行检验,如发现其中的次品数多于 1,就去调整设备. 以 X 表示一天中调整设备的次数,试求 $E(X)$. (设诸产品是否为次品是相互独立的.)

3. 有 3 只球,4 个盒子,盒子的编号为 1,2,3,4. 将球逐个独立地,随机地放入 4 个盒子中去. 以 X 表示其中至少有一只球的盒子的最小号码(例如 $X=3$ 表示第 1 号,第 2 号盒子是

空的，第 3 个盒子至少有一只球)，试求 $E(X)$.

4. (1) 设随机变量 X 的分布律为 $P\left\{X=(-1)^{j+1}\frac{3^j}{j}\right\}=\frac{2}{3^j}$，$j=1,2,\cdots$，说明 X 的数学期望不存在.

(2) 一盒中装有一只黑球，一只白球，作摸球游戏，规则如下：一次从盒中随机摸一只球，若摸到白球，则游戏结束，摸到黑球放回再放入一只黑球，然后再从盒中随机地摸一只球. 试说明要游戏结束的摸球次数 X 的数学期望不存在.

5. 设在某一规定的时间间隔里，某电气设备用于最大负荷的时间 X(以 min 计)是一个随机变量，其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{1\,500^2}x, & 0\leqslant x\leqslant 1\,500,\\ \dfrac{-1}{1\,500^2}(x-3\,000), & 1\,500<x\leqslant 3\,000,\\ 0, & \text{其他}.\end{cases}$$

求 $E(X)$.

6. (1) 设随机变量 X 的分布律为

X	-2	0	2
p_k	0.4	0.3	0.3

求 $E(X)$，$E(X^2)$，$E(3X^2+5)$.

(2) 设 $X\sim\pi(\lambda)$，求 $E[1/(X+1)]$.

7. (1) 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}e^{-x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求(i)$Y=2X$；(ii)$Y=e^{-2X}$的数学期望.

(2) 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且都服从(0,1)上的均匀分布(i)求 $U=\max\{X_1,X_2,\cdots,X_n\}$的数学期望，(ii)求 $V=\min\{X_1,X_2,\cdots,X_n\}$的数学期望.

8. 设随机变量(X,Y)的分布律为

Y \\ X	1	2	3
-1	0.2	0.1	0.0
0	0.1	0.0	0.3
1	0.1	0.1	0.1

(1) 求 $E(X)$，$E(Y)$.

(2) 设 $Z=Y/X$，求 $E(Z)$.

(3) 设$Z=(X-Y)^2$，求 $E(Z)$.

9. (1) 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}12y^2, & 0\leqslant y\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

求 $E(X),E(Y),E(XY),E(X^2+Y^2)$.

(2) 设随机变量 X,Y 的联合密度为

$$f(x,y)=\begin{cases}\dfrac{1}{y}\mathrm{e}^{-(y+x/y)}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

求 $E(X),E(Y),E(XY)$.

10. (1) 设随机变量 $X\sim N(0,1),Y\sim N(0,1)$ 且 X,Y 相互独立. 求 $E[X^2/(X^2+Y^2)]$.

(2) 一飞机进行空投物资作业,设目标点为原点 $O(0,0)$,物资着陆点为 (X,Y),X,Y 相互独立,且设 $X\sim N(0,\sigma^2),Y\sim N(0,\sigma^2)$,求原点到点 (X,Y) 间距离的数学期望.

11. 一工厂生产的某种设备的寿命 X(以年计)服从指数分布,概率密度为

$$f(x)=\begin{cases}\dfrac{1}{4}\mathrm{e}^{-x/4}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

工厂规定,出售的设备若在售出一年之内损坏可予以调换. 若工厂售出一台设备赢利 100 元,调换一台设备厂方需花费 300 元. 试求厂方出售一台设备净赢利的数学期望.

12. 某车间生产的圆盘直径在区间 (a,b) 服从均匀分布,试求圆盘面积的数学期望.

13. 设电压(以 V 计) $X\sim N(0,9)$. 将电压施加于一检波器,其输出电压为 $Y=5X^2$,求输出电压 Y 的均值.

14. 设随机变量 X_1,X_2 的概率密度分别为

$$f_1(x)=\begin{cases}2\mathrm{e}^{-2x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\qquad f_2(x)=\begin{cases}4\mathrm{e}^{-4x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(1) 求 $E(X_1+X_2),E(2X_1-3X_2^2)$.

(2) 又设 X_1,X_2 相互独立,求 $E(X_1X_2)$.

15. 将 n 只球($1\sim n$ 号)随机地放进 n 个盒子($1\sim n$ 号)中去,一个盒子装一只球. 若一只球装入与球同号的盒子中,称为一个配对. 记 X 为总的配对数,求 $E(X)$.

16. 若有 n 把看上去样子相同的钥匙,其中只有一把能打开门上的锁,用它们去试开门上的锁. 设取到每只钥匙是等可能的. 若每把钥匙试开一次后除去,试用下面两种方法求试开次数 X 的数学期望.

(1) 写出 X 的分布律.

(2) 不写出 X 的分布律.

17. 设 X 为随机变量,C 是常数,证明 $D(X)<E[(X-C)^2]$,对于 $C\neq E(X)$.(由于 $D(X)=E\{[X-E(X)]^2\}$,上式表明 $E[(X-C)^2]$ 当 $C=E(X)$ 时取到最小值.)

18. 设随机变量 X 服从瑞利分布,其概率密度为

$$f(x)=\begin{cases}\dfrac{x}{\sigma^2}\mathrm{e}^{-x^2/(2\sigma^2)}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 $\sigma>0$ 是常数. 求 $E(X),D(X)$.

19. 设随机变量 X 服从 Γ 分布，其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-x/\beta}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 $\alpha>0,\beta>0$ 是常数. 求 $E(X),D(X)$.

20. 设随机变量 X 服从几何分布，其分布律为

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

其中 $0<p<1$ 是常数. 求 $E(X),D(X)$.

21. 设长方形的高(以 m 计) $X\sim U(0,2)$，已知长方形的周长(以 m 计)为 20. 求长方形面积 A 的数学期望和方差.

22. (1) 设随机变量 X_1,X_2,X_3,X_4 相互独立，且有 $E(X_i)=i,D(X_i)=5-i,i=1,2,3,4$. 设 $Y=2X_1-X_2+3X_3-\frac{1}{2}X_4$. 求 $E(Y),D(Y)$.

(2) 设随机变量 X,Y 相互独立，且 $X\sim N(720,30^2)$，$Y\sim N(640,25^2)$，求 $Z_1=2X+Y,Z_2=X-Y$ 的分布，并求概率 $P\{X>Y\}$，$P\{X+Y>1\,400\}$.

23. 五家商店联营，它们每两周售出的某种农产品的数量(以 kg 计)分别为 X_1,X_2,X_3,X_4,X_5. 已知 $X_1\sim N(200,225)$，$X_2\sim N(240,240)$，$X_3\sim N(180,225)$，$X_4\sim N(260,265)$，$X_5\sim N(320,270)$，X_1,X_2,X_3,X_4,X_5 相互独立.

(1) 求五家商店两周的总销售量的均值和方差.

(2) 商店每隔两周进货一次，为了使新的供货到达前商店不会脱销的概率大于 0.99，问商店的仓库应至少储存多少千克该产品？

24. 卡车装运水泥，设每袋水泥重量 X(以 kg 计)服从 $N(50,2.5^2)$，问最多装多少袋水泥使总重量超过 2 000 的概率不大于 0.05.

25. 设随机变量 X,Y 相互独立，且都服从 $(0,1)$ 上的均匀分布.

(1) 求 $E(XY),E(X/Y),E[\ln(XY)],E[|Y-X|]$.

(2) 以 X,Y 为边长作一长方形，以 A,C 分别表示长方形的面积和周长，求 A 和 C 的相关系数.

26. (1) 设随机变量 X_1,X_2,X_3 相互独立，且有 $X_1\sim b(4,1/2)$，$X_2\sim b(6,1/3)$，$X_3\sim b(6,1/3)$，求 $P\{X_1=2,X_2=2,X_3=5\}$，$E(X_1X_2X_3)$，$E(X_1-X_2)$，$E(X_1-2X_2)$.

(2) 设 X,Y 是随机变量，且有 $E(X)=3,E(Y)=1,D(X)=4,D(Y)=9$，令 $Z=5X-Y+15$，分别在下列 3 种情况下求 $E(Z)$ 和 $D(Z)$.

(i) X,Y 相互独立，(ii) X,Y 不相关，(iii) X 与 Y 的相关系数为 0.25.

27. 下列各对随机变量 X 和 Y，问哪几对是相互独立的？哪几对是不相关的.

(1) $X\sim U(0,1),Y=X^2$.

(2) $X\sim U(-1,1),Y=X^2$.

(3) $X=\cos V,Y=\sin V,V\sim U(0,2\pi)$.

若 (X,Y) 的概率密度为 $f(x,y)$，

(4) $f(x,y)=\begin{cases}x+y, & 0<x<1,0<y<1,\\ 0, & \text{其他}.\end{cases}$

(5) $f(x,y)=\begin{cases}2y, & 0<x<1,0<y<1,\\ 0, & \text{其他}.\end{cases}$

28. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, & x^2+y^2\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

试验证X和Y是不相关的，但X和Y不是相互独立的.

29. 设随机变量(X,Y)的分布律为

Y \ X	-1	0	1
-1	$1/8$	$1/8$	$1/8$
0	$1/8$	0	$1/8$
1	$1/8$	$1/8$	$1/8$

验证X和Y是不相关的，但X和Y不是相互独立的.

30. 设A和B是试验E的两个事件，且$P(A)>0,P(B)>0$，并定义随机变量X,Y如下：

$$X=\begin{cases}1, & \text{若 }A\text{ 发生},\\ 0, & \text{若 }A\text{ 不发生},\end{cases}\qquad Y=\begin{cases}1, & \text{若 }B\text{ 发生},\\ 0, & \text{若 }B\text{ 不发生}.\end{cases}$$

证明若$\rho_{XY}=0$，则X和Y必定相互独立.

31. 设随机变量(X,Y)具有概率密度

$$f(x,y)=\begin{cases}1, & |y|<x,0<x<1,\\ 0, & \text{其他}.\end{cases}$$

求$E(X),E(Y),\mathrm{Cov}(X,Y)$.

32. 设随机变量(X,Y)具有概率密度

$$f(x,y)=\begin{cases}\dfrac{1}{8}(x+y), & 0\leqslant x\leqslant 2,0\leqslant y\leqslant 2,\\ 0, & \text{其他}.\end{cases}$$

求$E(X),E(Y),\mathrm{Cov}(X,Y),\rho_{XY},D(X+Y)$.

33. 设随机变量$X\sim N(\mu,\sigma^2),Y\sim N(\mu,\sigma^2)$，且设$X,Y$相互独立，试求$Z_1=\alpha X+\beta Y$和$Z_2=\alpha X-\beta Y$的相关系数(其中$\alpha,\beta$是不为零的常数).

34. (1) 设随机变量$W=(aX+3Y)^2,E(X)=E(Y)=0,D(X)=4,D(Y)=16,\rho_{XY}=-0.5$.求常数$a$使$E(W)$为最小，并求$E(W)$的最小值.

(2) 设随机变量(X,Y)服从二维正态分布，且有$D(X)=\sigma_X^2,D(Y)=\sigma_Y^2$.证明当$a^2=\sigma_X^2/\sigma_Y^2$时，随机变量$W=X-aY$与$V=X+aY$相互独立.

35. 设随机变量(X,Y)服从二维正态分布，且$X\sim N(0,3),Y\sim N(0,4)$，相关系数$\rho_{XY}=-1/4$，试写出$X$和$Y$的联合概率密度.

36. 已知正常男性成人血液中，每一毫升白细胞数平均是7 300，均方差是700.利用切比雪夫不等式估计每毫升含白细胞数在5 200～9 400之间的概率p.

37. 对于两个随机变量V,W，若$E(V^2),E(W^2)$存在，证明

$$[E(VW)]^2 \leqslant E(V^2)E(W^2). \tag{A}$$

这一不等式称为柯西一施瓦茨(Cauchy-Schwarz)不等式.

提示:考虑实变量 t 的函数

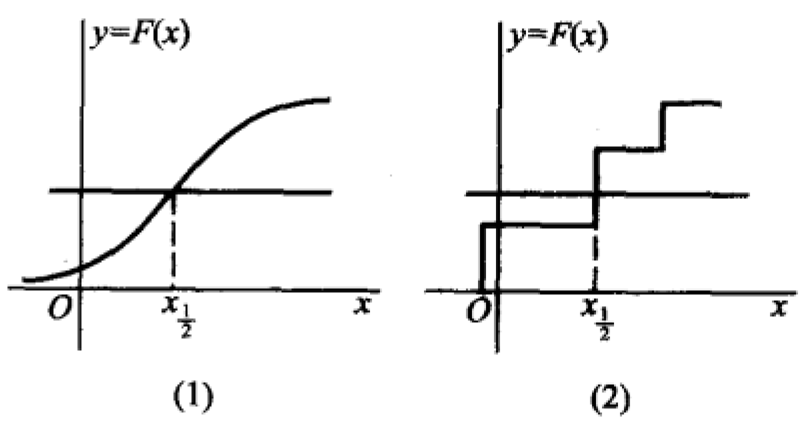
$$q(t)=E[(V+tW)^2]=E(V^2)+2tE(VW)+t^2E(W^2).$$

38. 中位数.

对于任意随机变量 X,满足以下两式

$$P\{X\leqslant x\}\geqslant\frac{1}{2},\qquad P\{X\geqslant x\}\leqslant\frac{1}{2}$$

的 x 称为 X 的**中位数**,记为 $x_{\frac{1}{2}}$ 或 M. 它是反映集中位置的一个数字特征,中位数总是存在的. 画出 X 的分布函数 $F(x)$的图. 如果 $F(x)$连续,那么 $x_{\frac{1}{2}}$ 是方程 $F(x)=\frac{1}{2}$ 的解(如题 38 图(1)),如果 $F(x)$有跳跃点(见题 38 图(2)),用垂直于横轴的线段联结后,得一连续曲线,它与直线 $y=1/2$ 的交点的横坐标即为 $x_{\frac{1}{2}}$.



题 38 图

(1) 设 X 的概率密度为

$$f(x)=\begin{cases}2e^{-2x}, & x\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

试求 X 的中位数 M.

(2) 设 X 服从柯西分布,其概率密度为

$$f(x)=\frac{b}{\pi[(x-a)^2+b^2]},b>0.$$

试求 X 的中位数 M.