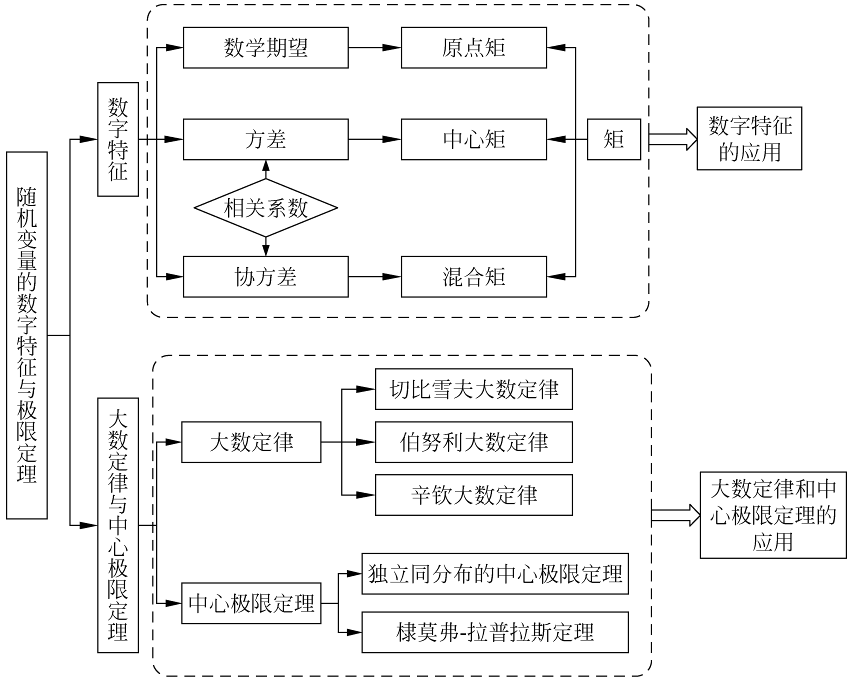
第3章

随机变量的数字特征与极限定理

3.1 知识网络图



3.2 精品课堂

第3章

随机变量的数字特征与极限定理

数学期望

EX 设某班40名学生的概率统计成绩及得分人数如下表所示。

分数	40	60	70	80	90	100
人数	1	6	9	15	7	2

则学生的平均成绩是总分÷总人数(分)，即

$$\frac{1\times40+6\times60+9\times70+15\times80+7\times90+2\times100}{1+6+9+15+7+2}=76.5(\text{分})$$

第3章 随机变量的数字特征与极限定理

离散型随机变量的数学期望

设离散型随机变量的分布律为

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

若级数 $\sum_k x_k p_k$ 绝对收敛，则称 $\sum_k x_k p_k$ 为 X 的数学期望，简称期望或均值，记作 $E(X)$，即

$$E(X)=\sum_k x_k p_k$$

一维随机变量的数字特征 1

请注意

(1) 定义中的绝对收敛是对可以取到无穷可数个值而言的。它保证级数的收敛性及级数的和与其中各项的次序无关。

(2) 随机变量的数学期望是一个实数，它是随机变量取值的概率平均。

(3) 当X服从某一分布时，也称其分布的数学期望为$E(X)$。

例1 甲乙两人练习射击打靶,命中环数分别记为X和Y,X与Y的分布律如下表所示,试评定他们的技术的优劣。

X	0	5	6	7	8	9	10
p	0.05	0.05	0.05	0.05	0.1	0.2	0.5

Y	0	5	6	7	8	9	10
p	0	0.01	0.09	0.15	0.35	0.35	0.05

解 为了评价甲乙技术的优劣,我们先来求随机变量和的平均值,即数学期望。

$$E(X)=0\times0.05+5\times0.05+6\times0.05+7\times0.05+8\times0.1+9\times0.2+10\times0.5=8.5$$

$$E(Y)=0\times0+5\times0.01+6\times0.09+7\times0.15+8\times0.35+9\times0.35+10\times0.05=8.09$$

从平均命中环数看,甲的射击水平比乙的高。

连续型随机变量的数学期望

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

$$E(X)=\int_{-\infty}^{+\infty}x\,f(x)\mathrm{d}x$$

为X的数学期望。

例2 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	8:30	8:50
到站时刻	9:10	9:30	9:50
概率分布律	1/6	3/6	2/6

(1)一旅客8:00到车站,求他候车时间的数学期望。

(2)一旅客8:20到车站,求他候车时间的数学期望。

例3 设随机变量X的概率密度为

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

求X的数学期望。

解 由连续型随机变量的数学期望计算公式得

$$E(X)=\int_{-\infty}^{+\infty}x\,f(x)\mathrm{d}x$$

$$=\int_{-1}^{0}x(1+x)\mathrm{d}x+\int_{0}^{1}x(1-x)\mathrm{d}x=0$$

解 设旅客的候车时间为X

(1) X的分布律为

X	10	30	50
p_k	1/6	3/6	2/6

$$E(X)=10\times\frac{1}{6}+30\times\frac{3}{6}+50\times\frac{2}{6}=33.33$$

(2) X的分布律为

X	10	30	50	70	90
p_k	3/6	2/6	1/6×1/6	3/6×1/6	2/6×1/6

$$E(X)=10\times\frac{3}{6}+30\times\frac{2}{6}+50\times\frac{1}{36}+70\times\frac{3}{36}+90\times\frac{2}{36}=27.22$$

随机变量的函数的数学期望

若随机变量$Y=g(X)$,且$E(Y)$存在,则有

$$E(Y)=Eg(X)\overset{\text{离散型}}{\underset{\text{连续型}}{=}}\begin{cases}\sum_i g(x_i)P\{X=x_i\}\\\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x\end{cases}$$

其中$f(x)$是连续型随机变量X的概率密度。

请注意

不是所有的随机变量都有数学期望的,例如随机变量X的分布律为

$$P\left\{X=(-1)^k\times\frac{2^k}{k}\right\}=\frac{1}{2^k},\quad k=1,2,\cdots$$

因为级数 $\sum_{k=1}^{\infty}(-1)^k\times\frac{2^k}{k}\times\frac{1}{2^k}=\sum_{k=1}^{\infty}(-1)^k\times\frac{1}{k}$

不绝对收敛,所以$E(X)$不存在。

常用结论

(1) $E(c)=c$,c为常数;

(2) $E(cX)=cE(X)$,c为常数;

(3) $E(X+c)=E(X)+c$;

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

(5) $E[g_1(X)+g_2(X)]=E[g_1(X)]+E[g_2(X)]$;

(6) 若X与Y相互独立,则$E(XY)=E(X)\,E(Y)$。

例4 一水果商每天卖出去的水果数量(单位：kg)服从参数为λ的泊松分布。如果每卖出1kg可得报酬a元，卖不掉就损失b元，若每天水果商批发进水果$n$kg，求其期望所得。

解 设水果商每日卖出的水果数量为Xkg，每日所得为$g(X)$，根据题意，得

$$g(X)=\begin{cases} an, & X\geqslant n \\ aX-b(n-X), & X<n \end{cases}$$

$$g(x)=\begin{cases} an, & x\geqslant n \\ (a+b)x-bn, & x<n \end{cases}$$

$$E[g(X)]=\sum_{k=0}^{\infty}g(k)\frac{\lambda^k}{k!}e^{-\lambda}$$

$$=\sum_{k=0}^{n-1}[(a+b)k-bn]\cdot\frac{\lambda^k}{k!}e^{-\lambda}+\sum_{k=n}^{\infty}an\frac{\lambda^k}{k!}e^{-\lambda}$$

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

$$=(a+b)\lambda e^{-\lambda}\sum_{k=0}^{n-2}\frac{\lambda^k}{k!}-n(a+b)e^{-\lambda}\sum_{k=0}^{n-1}\frac{\lambda^k}{k!}+na$$

例5 设随机变量X的分布律如表所示，求$E(X^2+X-1)$。

X	-1	0	1	2
p_k	0.1	0.2	0.3	0.4

解 $E(X^2+X-1)$

$=[(-1)^2+(-1)-1]\times 0.1$

$+[0^2+0-1]\times 0.2+[1^2+1-1]\times 0.3$

$+[2^2+2-1]\times 0.4=2.0$

例6 已知$X\sim N(0,1)$，求$E(X^2)$。

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

$$=\int_{-\infty}^{+\infty}-x\,d\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\right)$$

$$=-x\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$$=0+1=1$$

例7 一民航的载有20位旅客的送客车自机场开出，旅客有10个车站可以下车。如果到一个车站没有旅客下车就不停车。以X表示停车次数，求$E(X)$(设各旅客在10个站下车是等可能的且是否下车相互独立)。

解 令

$$X_i=\begin{cases} 0, & \text{在第}i\text{站无人下车} \\ 1, & \text{在第}i\text{站有人下车} \end{cases},\ i=1,2,\cdots,10$$

$$X=X_1+X_2+\cdots+X_{10}$$

$$P\{X_i=0\}=\left(\frac{9}{10}\right)^{20}$$

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

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

$$E(X)=E(X_1+X_2+\cdots+X_{10})$$

$$=E(X_1)+E(X_2)+\cdots+E(X_{10})$$

$$=10\left[1-\left(\frac{9}{10}\right)^{20}\right]=8.784$$

方差

方差是衡量随机变量取值波动*程度*的一个数字特征。

如何定义？

方差和标准差是表征随机变量取值分散或集中程度的数字特征。

如果随机变量$[X-E(X)]^2$的期望存在，则称

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

为随机变量X的方差，称$\sqrt{D(X)}$为X的标准差。

随机变量X的方差有以下计算公式。

$$D(X)=\begin{cases}\sum_i [x_i-E(X)]^2 P\{X=x_i\}, & \text{离散型}\\ \int_{-\infty}^{+\infty}[x_i-E(X)]^2 f(x)\mathrm{d}x, & \text{连续型}\end{cases}$$

请注意

方差反映的是X的取值与其数学期望之间的分散程度。

方差的常用结论

假设c为任意常数，下列各式中相应期望、方差都存在，则

(1) $D(c)=0$;

(2) $D(cX)=c^2D(X)$，$D(-X)=D(X)$;

(3) $D(X+c)=D(X)$;

(4) 若X和Y相互独立，则有$D(X\pm Y)=D(X)+D(Y)$;

(5) 若$X_1,X_2,\cdots,X_m$两两独立或两两不相关，则

$D(X_1,X_2,\cdots,X_m)=D(X_1)+D(X_2)+\cdots+D(X_m)$。

例8 设随机变量X具有概率密度

$$f(x)=\begin{cases}1+x\ , & -1\leqslant x<0\\ 1-x\ , & 0\leqslant x<1\\ 0\ , & \text{其他}\end{cases}$$，求$D(X)$。

解 因为 $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$

$$=\int_{-1}^{0}x(1+x)\mathrm{d}x+\int_{0}^{1}x(1-x)\mathrm{d}x=0$$

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

$$=\int_{-1}^{0}x^2(1+x)\mathrm{d}x+\int_{0}^{1}x^2(1-x)\mathrm{d}x=\frac{1}{6}$$

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

离散型随机变量

常见分布的期望、方差

分布	数学期望	方差
0-1分布	$E(X)=p$	$D(X)=pq$
二项分布	$E(X)=np$	$D(X)=npq$
泊松分布	$E(X)=\lambda$	$D(X)=\lambda$

连续型随机变量

常见分布的数学期望

分布	数学期望	方差
均匀分布	$E(X)=\frac{1}{2}(a+b)$	$D(X)=\frac{(b-a)^2}{12}$
指数分布	$E(X)=\frac{1}{\lambda}$	$D(X)=\frac{1}{\lambda^2}$
正态分布	$E(X)=\mu$	$D(X)=\sigma^2$
Γ分布	$E(X)=\frac{\alpha}{\beta}$	$D(X)=\frac{\alpha}{\beta^2}$

例9 设随机变量$X_1,X_2,\cdots,X_n(n>1)$相互独立且服从同一分布，其方差为$\sigma^2>0$。令$Y=\frac{1}{n}\sum_{i=1}^{n}X_i$，求$D(X_1+Y)$。

解 $D(X_1+Y)=D\left(\frac{1+n}{n}X_1+\frac{1}{n}X_2+\cdots+\frac{1}{n}X_n\right)$

$$=\frac{(1+n)^2}{n^2}D(X_1)+\frac{1}{n^2}D(X_2)+\cdots+\frac{1}{n^2}D(X_n)$$

$$=\frac{(1+n)^2}{n^2}\sigma^2+\frac{n-1}{n^2}\sigma^2$$

$$=\frac{n^2+3n}{n^2}\sigma^2=\frac{n+3}{n}\sigma^2$$

随机变量的矩

设X和Y是随机变量

(1) 若$E(X^k)\ (k=1,2,\cdots)$存在，则称它为X的k阶原点矩。

(2) 若$E\{[X-E(X)]^k\}\ (k=1,2,\cdots)$存在，则称它为$X$的$k$阶中心矩。

二维随机变量的数字特征 2

协方差

设(X,Y)是二维随机变量，如果X,Y的期望$E(X)$，$E(Y)$存在，且$E\{[X-E(X)][Y-E(Y)]\}$也存在，则称$[X-E(X)][Y-E(Y)]$的数学期望为两个随机变量X与Y的协方差，记作$\mathrm{Cov}(X,Y)$，即

$$\begin{aligned}\mathrm{Cov}(X,Y)&=E\{[X-E(X)][Y-E(Y)]\}\\&=E(XY)-E(X)E(Y)\end{aligned}$$

其中 $E(XY)=\begin{cases}\sum\limits_i\sum\limits_j x_iy_jP\{X=x_i,Y=y_j\}, & \text{离散型}\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y, & \text{连续型}\end{cases}$

4个等价结论

收藏夹

(1) $\mathrm{Cov}(X,Y)=0$;

(2) X与Y不相关;

(3) $E(XY)=E(X)E(Y)$;

(4) $D(X+Y)=D(X)+D(Y)$。

协方差的性质

(1) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$;

(2) $\mathrm{Cov}(X,X)=D(X)$;

(3) $\mathrm{Cov}(X,c)=0$;

(4) $\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y)$;

(5) $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$;

(6) $[\mathrm{Cov}(X,Y)]^2\leqslant D(X)D(Y)$;

(7) 如果X与Y独立，则$\mathrm{Cov}(X,Y)=0$;

(8) 对任意两个随机变量X，Y，如果其方差存在，则$X\pm Y$的方差也存在，且

$D(X\pm Y)=D(X)\pm 2\mathrm{Cov}(X,Y)+D(Y)$

例10 设二维连续随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{4}x(1+3y^2), & 0<x<2,0<y<1\\ 0, & \text{其他}\end{cases},$$

求 $E(X)$，$E(Y)$。

解 $E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y$

$=\int_0^2 x\cdot\frac{1}{4}x\mathrm{d}x\int_0^1(1+3y^2)\mathrm{d}y=\frac{4}{3}$

$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y$

$=\int_0^2\frac{1}{4}x\mathrm{d}x\int_0^1 y(1+3y^2)\mathrm{d}y=\frac{5}{8}$

相关系数

如果X，Y的方差都不等于零，则称

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

为随机变量X与Y的相关系数。

例11 对于例10中的二维连续随机变量(X,Y)，求$E(X+Y)$，$E(XY)$。

解 $E(X+Y)=E(X)+E(Y)$

$=\frac{4}{3}+\frac{5}{8}=\frac{47}{24}$

可以验证X,Y相互独立，所以

$E(XY)=E(X)E(Y)=\frac{4}{3}\times\frac{5}{8}=\frac{5}{6}$

相关系数的性质

(1) $\rho_{XY}=\rho_{YX}$;

(2) $\rho_{XX}=1$;

(3) $|\rho_{XY}|\leqslant 1$;

(4) 若$X=aY+b$，则当$a>0$时，$\rho_{XY}=1$；当$a<0$时，$\rho_{XY}=-1$;

(5) $\rho_{XY}=1\Leftrightarrow X$和$Y$(以概率为1)互为线性函数;

(6) $\rho_{XY}=0$，则称X与Y不相关。

例12 设(X,Y)服从二维正态分布$N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$，求$\mathrm{Cov}(X,Y)$和$\rho_{XY}$。

解 因为$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，因而

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

令 $u=\frac{x-\mu_1}{\sigma_1}$，$u=\frac{y-\mu_2}{\sigma_2}$，则

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)$$

$$e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}\mathrm{d}x\mathrm{d}y$$

令 $u=\frac{x-\mu_1}{\sigma_1}$, $v=\frac{y-\mu_2}{\sigma_2}$, $\mathrm{d}u\mathrm{d}v=\frac{1}{\sigma_1\sigma_2}\mathrm{d}x\mathrm{d}y$, 则

$$\mathrm{Cov}(X,Y)=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}uv\cdot e^{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)}\mathrm{d}u\mathrm{d}v$$

$$=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}uv\cdot e^{-\frac{1}{2(1-\rho^2)}[(u^2-\rho v)^2+(1-\rho^2)v^2]}\mathrm{d}u\right]\mathrm{d}v$$

$$=\frac{\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left[ve^{-\frac{1}{2}v^2}\frac{1}{\sqrt{2\pi}\cdot\sqrt{1-\rho^2}}\times\int_{-\infty}^{+\infty}ue^{-\frac{1}{2(1-\rho^2)}(u-\rho v)^2}\mathrm{d}u\right]\mathrm{d}v$$

$$=\frac{\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\rho v^2e^{-\frac{1}{2}v^2}\mathrm{d}v=\sigma_1\sigma_2\rho$$

注意 考虑上面计算过程中的最后两步积分

$\frac{1}{\sqrt{2\pi}\cdot\sqrt{1-\rho^2}}\times\int_{-\infty}^{+\infty}ue^{-\frac{1}{2(1-\rho^2)}(u-\rho v)^2}\mathrm{d}u$ 是正态分布 $N(\rho v,1-\rho^2)$ 的数学期望，因而等于 ρv。

又设 $X_1\sim N(0,1)$,

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}v^2e^{-\frac{1}{2}v^2}\mathrm{d}v=E(X_1^2)=D(X_1)+E^2(X_1)=1$$

$$\rho_{XY}=\frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{\sigma_1\sigma_2\rho}{\sigma_1\sigma_2}=\rho$$

二维正态分布中的第五个参数

独立性与相关性

(1) 若随机变量X与Y独立，则X与Y一定不相关；
但若X与Y不相关，则X与Y可能独立，也可能不独立。

(2) 若随机变量X与Y的联合分布是二维正态分布，
则X与Y独立 $\Leftrightarrow$ X与Y不相关。

例13 设(X,Y)服从 $D=\{(x,y)\,|\,x^2+y^2\leqslant 1\}$ 上的均匀分布，求$\mathrm{Cov}(X,Y)$和ρ_{XY}，并且讨论X和Y的独立性。

解 D是以原点为圆心，1为半径的圆，其面积等于π，故(X,Y)的密度函数为

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

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\frac{1}{\pi}\iint_{x^2+y^2<1}x\mathrm{d}x\mathrm{d}y=0$$

同理 $E(Y)=0$。

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{\pi}\iint_{x^2+y^2<1}xy\mathrm{d}x\mathrm{d}y=0$$

又当$|x|\leqslant 1$时，$f_X(x)=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y=\frac{2}{\pi}\sqrt{1-x^2}$

又当$|y|\leqslant 1$时，$f_Y(y)=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}\mathrm{d}x=\frac{2}{\pi}\sqrt{1-y^2}$

显然 $f_X(x)f_Y(y)\neq f(x,y)$，故X和Y不是相互独立的，

故 $\rho_{XY}=0$ 不是X,Y相互独立的充分条件。

例14 已知离散型随机变量X的分布律为

X	-2	-1	0	2	1
p_k	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

令 $Y=X^2$，问X和Y是否不相关？
X和Y是否相互独立？

解 可以求得

$$E(X)=-2\times\frac{1}{5}+(-1)\times\frac{1}{5}+0\times\frac{1}{5}+1\times\frac{1}{5}+2\times\frac{1}{5}=0$$

$$E(Y)=E(X^2)=(-2)^2\times\frac{1}{5}+(-1)^2\times\frac{1}{5}+0^2\times\frac{1}{5}+1^2\times\frac{1}{5}+2^2\times\frac{1}{5}=2$$

$$E(XY)=E(X^3)=(-2)^3\times\frac{1}{5}+(-1)^3\times\frac{1}{5}+0^3\times\frac{1}{5}+1^3\times\frac{1}{5}+2^3\times\frac{1}{5}=0$$

$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0\Rightarrow$ X和Y不相关

X和Y之间有函数关系 $Y=X^2$ $\Rightarrow$ X和Y之间不独立

例15 设 X 和 Y 的联合概率密度函数为

$$f(x,y)=\begin{cases}2-x-y & ,0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\\ 0 & ,\quad 其他\end{cases}$$

讨论 X 和 Y 的独立性和相关性。

解 先求两个边缘概率密度

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_0^1(2-x-y)\mathrm{d}y=\frac{3}{2}-x & ,0\leqslant x\leqslant 1\\ 0 & ,其他\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_0^1(2-x-y)\mathrm{d}x=\frac{3}{2}-y & ,0\leqslant y\leqslant 1\\ 0 & ,其他\end{cases}$$

$$f_X(x)f_Y(y)=\begin{cases}\left(\frac{3}{2}-x\right)\left(\frac{3}{2}-y\right) & ,0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\\ 0 & ,\quad 其他\end{cases}\neq f(x,y)$$

所以 X 和 Y 不相互独立。

$$E(X)=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x=\int_0^1x\left(\frac{3}{2}-x\right)\mathrm{d}x=\frac{5}{12}$$

$$E(Y)=\int_{-\infty}^{+\infty}yf_Y(y)\mathrm{d}y=\int_0^1y\left(\frac{3}{2}-y\right)\mathrm{d}y=\frac{5}{12}$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^1xy(2-x-y)\mathrm{d}y=\frac{1}{6}$$

$$\rho_{XY}=\frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)D(Y)}}=\frac{\frac{1}{6}-\frac{5}{12}\times\frac{5}{12}}{\sqrt{D(X)D(Y)}}\neq 0$$

所以 X 和 Y 相关。

原点矩和中心矩

定义 X 和 Y 是随机变量

(1) 若 $E(X^k)$ $(k=1,2,\cdots)$ 存在，则称它为 X 的 k 阶原点矩。

(2) 若 $E\left\{\left[X-E(X)\right]^k\right\}(k=1,2,\cdots)$ 存在，则称它为 X 的 k 阶中心矩。

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

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

矩的计算

(1) 原点矩的计算公式为

$$E(X^k)=\begin{cases}\sum_i x_i^kP\{X=x_i\} & (离散型)\\ \int_{-\infty}^{+\infty}x^kf(x)\mathrm{d}x & (连续型)\end{cases}$$

(2) 常用的 k 阶中心矩和 k 阶原点矩的换算公式。

设 $A_k=E(X^k)$, $B_k=E\left\{\left[X-E(X)\right]^k\right\}$,

数学期望是一阶原点矩 A_1，

方差是二阶中心矩 B_2，

原点矩和中心矩前4阶之间的转换公式为

$$\begin{cases}A_1=E(X)\\ A_2=B_2+A_1^2\\ A_3=B_3+3A_1B_2+A_1^3\\ A_4=B_4+4A_1B_3+6A_1^2B_2+A_1^4\end{cases}\qquad\begin{cases}B_1=0\\ B_2=A_2-A_1^2\\ B_3=A_3-3A_1A_2+2A_1^3\\ B_4=A_4-4A_1A_3+6A_1^2A_2-3A_1^4\end{cases}$$

例16 设随机变量 X 的概率分布律为

X	0	1	2	3	4	5
p_k	0	0.05	0.20	0.35	0.28	0.12

求随机变量 X 的4阶原点矩和3阶中心矩。

解 X 的4阶原点矩为

$$\begin{aligned}E(X^4)&=0^4\times 0+1^4\times 0.05+2^4\times 0.20\\&\quad+3^4\times 0.35+4^4\times 0.28+5^4\times 0.12\\&=178.28\end{aligned}$$

因为 $E(X)=0\times0+1\times0.05+2\times0.20+3\times0.35$
$+4\times0.28+5\times0.12$
$=3.22$

所以X的3阶中心矩为

$$E\{[X-E(X)]^3\}=E[(X-3.22)^3]$$

$$=(1-3.22)^3\times0.05+(2-3.22)^3\times0.20+(3-3.22)^3\times0.35+$$
$$+(4-3.22)^3\times0.28+(5-3.22)^3\times0.12$$
$$=-0.104\,304$$

例17 设随机变量X在区间(a,b)上服从均匀分布，求其k阶原点矩和k阶中心矩。

解 因为 $E(X)=\dfrac{a+b}{2}$，对于 $k=1,2,\cdots$

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

$$E\{[X-E(X)]^k\}=\frac{1}{b-a}\int_a^b[x-E(X)]^k\mathrm{d}x=\frac{1}{b-a}\int_a^b\left(x-\frac{b+a}{2}\right)^k\mathrm{d}x$$

$$=\frac{1}{(k+1)(b-a)}\left[\left(b-\frac{b+a}{2}\right)^{k+1}-\left(a-\frac{b+a}{2}\right)^{k+1}\right]$$

$$=\frac{1}{2^{k+1}(k+1)(b-a)}[(b-a)^{k+1}-(-1)^{k+1}\times(b-a)]$$

$$=\begin{cases}\dfrac{(b-a)^k}{2^k(k+1)}, & k\text{为偶数}\\ 0, & k\text{为奇数}\end{cases}$$

协方差矩阵

设二维随机变量(X_1,X_2)的关于X_1和X_2的二阶中心矩及二阶中心混合矩

$$C_{ij}=E\{[X_i-E(X_i)][X_j-E(X_j)]\},\ i,j=1,2$$

都存在，则称矩阵

$$\mathrm{C}=\begin{pmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{pmatrix}$$

为二维随机变量(X_1,X_2)的协方差矩阵。

例18 设二维随机变量 $(X_1,X_2)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$ 写出(X_1,X_2)的协方差矩阵。

解 因为 $X_1\sim N(\mu_1,\sigma_1^2)$，$X_2\sim N(\mu_2,\sigma_2^2)$

因此 $E(X_1)=\mu_1,E(X_2)=\mu_2$

$$C_{11}=D(X_1)=E\{[X_1-E(X_1)]^2\}=\sigma_1^2$$

$$C_{22}=D(X_2)=E\{[X_2-E(X_2)]^2\}=\sigma_2^2$$

$$C_{12}=C_{21}=\mathrm{Cov}(X_1,X_2)=\mathrm{Cov}(X_2,X_1)$$
$$=E\{[X_1-E(X_1)][X_2-E(X_2)]\}=\rho\sigma_1\sigma_2$$

随机变量(X_1,X_2)的协方差矩阵为 $C=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$

例19 设二维随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{6}{7}\left(x^2+\dfrac{1}{2}xy\right), & 0<x<1,0<y<2\\ 0, & \text{其他}\end{cases}$$

求(X,Y)的协方差矩阵和相关系数。

解 $E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\frac{6}{7}\int_0^1x\mathrm{d}x\int_0^2\left(x^2+\frac{1}{2}xy\right)\mathrm{d}y$

$$=\frac{6}{7}\int_0^1(2x^3+x^2)\mathrm{d}x=\frac{5}{7}$$

$$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=\frac{6}{7}\int_0^2y\mathrm{d}y\int_0^1\left(x^2+\frac{1}{2}xy\right)\mathrm{d}x$$
$$=\frac{6}{7}\int_0^2\left(\frac{1}{3}y+\frac{1}{4}y^2\right)\mathrm{d}x=\frac{8}{7}$$

$$E(X^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^2f(x,y)\mathrm{d}x\mathrm{d}y=\frac{6}{7}\int_0^1x^2\mathrm{d}x\int_0^2\left(x^2+\frac{1}{2}xy\right)\mathrm{d}y=\frac{39}{70}$$

$$D(X)=E(X^2)-E^2(X)=\frac{39}{70}-\frac{25}{49}=\frac{23}{490}$$

$$E(Y^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y^2f(x,y)\mathrm{d}x\mathrm{d}y=\frac{6}{7}\int_0^2y^2\mathrm{d}y\int_0^1\left(x^2+\frac{1}{2}xy\right)\mathrm{d}x=\frac{34}{21}$$

$$D(Y)=E(Y^2)-E^2(Y)=\frac{34}{21}-\frac{64}{49}=\frac{46}{147}$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\frac{6}{7}\int_0^2y\mathrm{d}y\int_0^1x\left(x^2+\frac{1}{2}xy\right)\mathrm{d}x=\frac{17}{21}$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{17}{21}-\frac{5}{7}\times\frac{8}{7}=-\frac{1}{147}$$

所以得到 (X,Y)的协方差矩阵为

$$\begin{bmatrix}\dfrac{23}{490} & -\dfrac{1}{147}\\ -\dfrac{1}{147} & \dfrac{46}{147}\end{bmatrix}$$

X与Y的相关系数为

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{-\dfrac{1}{147}}{\sqrt{\dfrac{23}{490}\times\dfrac{46}{147}}}=-\frac{\sqrt{15}}{69}$$

协方差矩阵性质

收藏夹

(1) 协方差矩阵为对称矩阵。

(2) $C_{ii}=D(X_i)$，$i=1,2,\cdots,n$。

(3) $C_{ij}{}^2\leqslant C_{ii}\cdot C_{jj}$，$i,j=1,2,\cdots,n$。

(4) C是非负定的，即对任意的n维向量 $\alpha=(\alpha_1,\alpha_2,\cdots\alpha_n)^{\mathrm{T}}$，$\alpha^{\mathrm{T}}C\alpha\geqslant 0$。

极限定理最重要的两种

1 大数定律

2 中心极限定理

思考与练习！

(1) 设随机变量X在区间$[-3,3]$上服从均匀分布，求X的各阶原点矩和中心矩。

(2) 设随机变量X的分布律为

X	-2	-1	0	1	2
p_k	0.2	0.2	0.1	0.2	0.3

求随机变量X的三阶原点矩和三阶中心矩。

切比雪夫不等式

定理 设随机变量X的期望$E(X)$与方差$D(X)$存在，则对任意的$\varepsilon>0$，有

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2}$$

请注意 (1) 在上式中取$\varepsilon=k\sqrt{D(X)}$，则有

$$P\left\{|X-E(X)|\geqslant k\sqrt{D(X)}\right\}\leqslant\frac{1}{k^2}。$$

(2) 切比雪夫不等式的另外一种形式为

$$P\{|X-E(X)|<\varepsilon\}\geqslant 1-\frac{D(X)}{\varepsilon^2}。$$

(3) 设有随机变量$X_1,X_2,\cdots,X_n$，其中$D(X_i)=\sigma_i^2$，$\rho_{X_iX_j}=\rho$，$i,j=1,2,\cdots,n,i\neq j$，写出$X_1,X_2,\cdots,X_n$的协方差矩阵。

(4) 设随机变量X和Y在圆域$x^2+y^2\leqslant a^2$上服从均匀分布，求X和Y的协方差矩阵。

切比雪夫不等式的作用

不等式给出了在未知X的分布的情况下，对概率

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

或 $P\{|X-E(X)|<\varepsilon\}$

的一种估计。

极限定理 3

例20 已知正常成人男性的血液中，每一毫升血液中白细胞数平均是7300，均方差为700。利用切比雪夫不等式估计每毫升血液中白细胞数为5200~9400的概率。

解 设每毫升血液中白细胞数为X，

依题意，$E(X)=7300$，$D(X)=700^2$

所求概率为

$$P\{5200\leqslant X\leqslant 9400\}$$

$$=P\{5200-7300\leqslant X-7300\leqslant 9400-7300\}$$

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

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

由切比雪夫不等式

$$P\{|X - E(X)| \leqslant 2100\} \geqslant 1 - \frac{D(X)}{(2100)^2} = 1 - \left(\frac{700}{2100}\right)^2 = 1 - \frac{1}{9} = \frac{8}{9}$$

即估计每毫升血液中白细胞数为5200~9400的概率不小于$\frac{8}{9}$。

定义 设$\{X_n\}$为一随机变量序列，X_n的数学期望$E(X_n)$存在，$n=1,2,\cdots$。令$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$

如果对任意的$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\{|\bar{X}_n - E(\bar{X}_n)| < \varepsilon\} = 1$$

则称$\{X_n\}$服从大数定律。

满足切比雪夫定理条件的随机变量序列$\{X_n\}$服从大数定律。

几个常用的大数定律

定义 设$\{X_n\}$ $(n=1,2,\cdots)$是随机变量序列，若对每一个$k \geqslant 1$, $X_1, X_2, \cdots, X_k$都是相互独立的，则称序列$\{X_n\}$是相互独立的。

定义 设$\{X_n\}$ $(n=1,2,\cdots)$是随机变量序列，a是一个常数，若对于任意的$\varepsilon > 0$有

$$\lim_{n\to\infty} P\{|X_n - a| < \varepsilon\} = 1$$

则称随机变量序列$\{X_n\}$依概率收敛于a，记成

$$X_n \xrightarrow{P} a$$

定理 伯努利(Bernoulli)定理。

设n_A是n次重复独立试验中事件A发生的次数，P是事件A在每次试验中发生的概率，则对于任意的$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1$$

伯努利定理说明

随着试验次数的增加，事件发生的频率逐渐稳定于事件发生的概率。

切比雪夫定理

设随机变量$X_i (i=1,2,\cdots)$相互独立，且有相同的有限的数学期望和方差。

$$E(X_i) = \mu,\ D(X_i) = \sigma^2,\ i = 1,2,\cdots$$

作前n个变量的算术平均$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$。

则对任意的$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1$$

伯努利定理作用

- 为用事件的频率去估计事件的概率提供了理论依据。
- 大数定理都要求随机变量具有方差，或方差有界，当随机变量序列具有相同的分布时，这一要求可以去掉。

思考

如果n个随机变量是相互独立的，且具有相同的有限的数学期望和方差，当n很大时，这n个随机变量的算术平均值以很大的概率接近它们的数学期望。

切比雪夫定理作用

定理给出了平均值稳定性的科学描述。

定理 辛钦(Khintchine)大数定理。

设随机变量序列$X_i (i=1,2,\cdots)$满足相互独立且有相同的分布，且具有有限的数学期望

$$E(X_i) = \mu\ (i = 1, 2, \cdots)$$

则对于任意的$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1$$

请注意

（1）伯努利大数定理是辛钦大数定理的特殊情况。

（2）当 n 很大时，随机变量在 n 次观测中的算术平均值 $\bar{X}$ 具有稳定性，它依概率收敛于其数学期望。

辛钦大数定理作用

为我们寻求随机变量的数学期望提供了一条切实可行的方法。

说明

令 $Y_n=\dfrac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}$，由上述定理，当 $n\to\infty$ 时，Y_n 的分布函数 $F_n(x)$ 依概率收敛于标准正态分布的分布函数 $\Phi(x)$。因而当 n 很大时，Y_n 近似服从标准正态分布。即

$$P\left\{\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}\approx\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt=\Phi(x)$$

中心极限定理的背景

实例 炮弹射击的落点与目标的偏差，就受许多随机因素的影响。

- 如瞄准时的误差。
- 空气阻力所产生的误差。
- 炮弹或炮身结构所引起的误差等。
- 对我们来说重要的是这些随机因素的总影响。

结论

当 n 很大时，$\sum_{i=1}^{n}X_i$ 近似服从正态分布 $N(n\mu,n\sigma^2)$，只要 n 很大，就能通过正态分布的分布函数值给出 $\sum_{i=1}^{n}X_i$ 的分布函数的近似值。

结论

大量的试验观察表明，这些随机变量之和的分布近似服从正态分布，它们是正态分布的源泉。这也是我们为什么总在强调正态分布在概率统计中的重要性的原因之一。中心极限定理从理论上证明了这一事实。

定理 棣莫弗-拉普拉斯(DeMoivre - Laplace)定理。

设随机变量 $X_n(n=1,2,\cdots)$ 服从 $B(n,p)$，则对任意实数 x，恒有

$$\lim_{n\to\infty}P\left\{\frac{X_n-np}{\sqrt{npq}}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt$$

中心极限定理

定理 列维-林德伯格(Levy-Lindberg)定理。

设随机变量 $X_k(k=1,2,\cdots)$ 独立同分布，且具有有限的数学期望和方差。

$$E(X_k)=\mu,\ D(X_k)=\sigma^2\neq 0,\ k=1,2,\cdots$$

则对任意的 $x\in(-\infty,+\infty)$，

$$\lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$$

二项分布的又一个近似计算公式

$X\sim B(n,p)$ 当 n 很大时，对任意的 $a<b$ 有

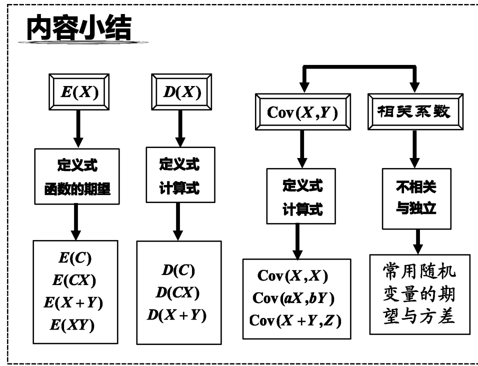
$$P\left\{a<\frac{X_n-np}{\sqrt{npq}}\leqslant b\right\}\approx\Phi(b)-\Phi(a)$$

其中 $\Phi(x)$ 是标准正态分布的分布函数。

请注意 (1) DeMoivre - Laplace定理说明，当 n 充分大时，$\dfrac{X_n-np}{\sqrt{npq}}$ 近似服从标准正态分布，而 n 较小时，这种近似不能用。

(2) 在我们作的和 $\sum_{i=1}^{n} X_i$ 中,每一个随机变量都为总和的变化提供了一个不可忽视的量,而每一个单独的项都不可能给总和做出很大的贡献。

(3) 由中心极限定理我们可以看出,总和虽然近似服从正态分布,但被加项不必服从正态分布。



3.3　达标实训

数字特征表现了随机变量某一个侧面的统计规律性,不需要像分布函数、分布律、概率密度一样去全面考查随机变量的整个变化情况。通过本节达标实训的训练,希望读者达到以下要求。

一级达标要求:深刻理解数字特征的基本概念;掌握随机变量的数学期望与方差的性质;会计算随机变量的期望与方差;熟记常见随机变量的期望与方差公式;掌握原点矩与中心矩的计算公式。

在此基础之上,实现二级达标要求:掌握协方差及相关系数的概念与性质;会计算随机变量的函数的数学期望;理解协方差矩阵与混合矩;会用切比雪夫不等式估计事件的概率;理解大数定律与中心极限定理。

"一级"和"二级"达标要求是学习本章必须达到的要求。

达标实训 Ⅰ

3-1　设随机变量 X 的分布律为

X	-2	0	2
p_k	0.4	0.3	0.3

求 $E(X)$,$E(X^2)$,$E(3X^2+5)$。

【解】 $E(X)=-2\times0.4+0\times0.3+2\times0.3=-0.2$

$E(X^2)=(-2)^2\times0.4+(0)^2\times0.03+2^2\times0.3=2.8$

$E(3X^2+5)=E(3X^2)+E(5)=3E(X^2)+5=13.4$

3-2 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-4\\ 0.35, & -4\leqslant x<1\\ 0.85, & 1\leqslant x<4\\ 1, & x\geqslant 4\end{cases}$$

求 $E(X)$。

【解】 由分布函数可知,X 为离散型随机变量。

$$P\{X=-4\}=0.35,\quad P\{X=1\}=0.5,\quad P\{X=4\}=0.15$$

$$E(X)=0.35\times(-4)+0.5\times1+0.15\times4=-0.3$$

3-3 连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases}kx^a, & 0<x<1\\ 0, & \text{其他}\end{cases}$$

又知 $E(X)=0.75$,求 k,a 的值。

【解】 $E(X)=\int_0^1 kx^{a+1}\mathrm{d}x=\frac{k}{a+2}x^{a+2}\Big|_0^1=\frac{k}{a+2}=0.75$

$$\int_0^1 kx^a\mathrm{d}x=\frac{k}{a+2}x^{a+1}\Big|_0^1=\frac{k}{a+1}=1$$

解得 $k=3$,$a=2$。

3-4 设随机变量 X 服从参数为 λ 的泊松分布,且已知 $E[(X-1)(X-2)]=1$,试求 λ。

【解】 $E[(X-1)(X-2)]=E(X^2-3X+2)$

$E(X^2-3X+2)=E(X^2)-3E(X)+E(2)=1$

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

已知 X 服从参数 λ 的泊松分布,故

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

$\because$ $E[(X-1)(X-2)]=1$ $\quad\therefore(\lambda-1)^2=0$ 得 $\lambda=1$

3-5 设离散型随机变量 X 的分布律为

X	-2	0	1
P	$3/10$	$1/5$	$1/2$

求 $E(2X^2-1)$,$D(2X^2-1)$。

【解】 $2X^2-1$ 的所有取值为 $-1,1,7$,

可得

$2X^2-1$	-1	1	7
P	$\frac{1}{5}$	$\frac{1}{2}$	$\frac{3}{10}$

$\therefore$ $E(2X^2-1)=(-1)\times\frac{1}{5}+1\times\frac{1}{2}+7\times\frac{3}{10}=\frac{12}{5}$

$$D(2X^2-1)=\left[(-1)-\frac{12}{5}\right]^2\times\frac{1}{5}+\left(1-\frac{12}{5}\right)^2\times\frac{1}{2}+\left(7-\frac{12}{5}\right)^2\times\frac{3}{10}=\frac{241}{25}$$

3-6　设随机变量 X 的概率密度为

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

求 $D(X)$。

【解】　$E(X)=\int_0^1 x^2\mathrm{d}x+\int_1^2 x(2-x)\mathrm{d}x=\frac{1}{3}+\frac{2}{3}=1$

$$E(X^2)=\int_0^1 x^3\mathrm{d}x+\int_1^2 x^2(2-x)\mathrm{d}x=\frac{1}{4}+\frac{11}{12}=\frac{7}{6}$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{7}{6}-1=\frac{1}{6}$$

3-7　设随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}0, & x\leqslant 0\\ \dfrac{3}{(x+1)^4}, & x>0\end{cases}$$

求 $E(X)$,$D(X)$。

【解】　$E(X)=\int_0^{+\infty} x\,\frac{3}{(x+1)^4}\mathrm{d}x=\frac{1}{2}$

$$E(X^2)=\int_0^{+\infty} x^2\,\frac{3}{(x+1)^4}\mathrm{d}x=1$$

$\therefore$ $D(X)=E(X^2)-[E(X)]^2=1-\frac{1}{4}=\frac{3}{4}$

3-8　设 X 的概率密度为 $f(x)=\begin{cases}|x|, & |x|<1\\ 0, & |x|\geqslant 1\end{cases}$,求 X 的数学期望和方差。

【解】　$E(X)=\int_{-1}^1 |x|x\mathrm{d}x=0$

$$E(X^2)=\int_{-1}^1 x^2|x|\mathrm{d}x=2\int_0^1 x^2x\mathrm{d}x=\frac{1}{2}$$

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

3-9　设 X 的概率密度为 $f(x)=\begin{cases}3x^2, & 0\leqslant x<1\\ 0, & \text{其他}\end{cases}$,求 X 的数学期望和方差。

【解】 $E(X)=\int_0^1 x\cdot 3x^2\mathrm{d}x=\frac{3}{4}$

$$D(X)=E(X^2)-[E(X)]^2=\int_0^1 x^2\cdot 3x^2\mathrm{d}x-\left(\frac{3}{4}\right)^2=\frac{3}{80}$$

3-10 设随机变量 X 的概率密度为

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

求 X 的期望值和方差。

【解】 根据概率密度的性质可得 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$,即

$$\int_0^2 Ax^2(x-2)^2\mathrm{d}x=1$$

解得 $A=\frac{15}{16}$

$$E(X)=\frac{15}{16}\int_0^2 x\cdot x^2(x-2)^2\mathrm{d}x=1$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{15}{16}\int_0^1 x\cdot x^2(x-2)\mathrm{d}x-[E(X)]^2=\frac{1}{7}$$

3-11 设二维连续随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases}\frac{1}{4}x(1+3y^2), & 0<x<2,0<y<1\\ 0, & \text{其他}\end{cases}$$

求 $E(X)$,$E(Y)$。

【解】

$$E(X)=\int_0^1\int_0^2\frac{1}{4}x^2(1+3y^2)\mathrm{d}x\mathrm{d}y=\int_0^1(1+3y^2)\mathrm{d}y\int_0^2\frac{1}{4}x^2\mathrm{d}x=\frac{4}{3}$$

$$E(Y)=\int_0^2\int_0^1\frac{1}{4}xy(1+3y^2)\mathrm{d}y\mathrm{d}x=\frac{5}{8}$$

3-12 设随机变量 X 在区间$[a,b]$上服从均匀分布,已知 $E(X)=1$,$E(X^2)=4$,求 a,b。

【解】 由题可知 $\begin{cases}E(X)=\frac{a+b}{2}=1\\ D(X)=\frac{(b-a)^2}{12}=E(X^2)-[E(X)]^2=3\end{cases}$

$\therefore\begin{cases}a=-2\\ b=4\end{cases}$ 或 $\begin{cases}a=4\\ b=-2\end{cases}$ 又$\because$区间(a,b) $\therefore\begin{cases}a=-2\\ b=4\end{cases}$

3-13 已知随机变量 X 服从二项分布,$E(X)=12$,$D(X)=8$,求 p 和 n。

【解】 依据二项分布的参数与期望和方差之间的关系,可得

$$\begin{cases}E(X)=np=12\\ D(X)=np(1-p)=8\end{cases}\quad\therefore\begin{cases}n=36\\ p=\frac{1}{3}\end{cases}$$

3-14 设一次试验成功的概率为 p，进行 120 次独立重复试验，问当 p 为何值时成功次数的标准差最大？并求标准差的最大值。

【解】 这是一个 120 重的伯努利模型，$X \sim B(120, p)$，故 $D(X)=120p(1-p)$，$D'(X)=120-240p$，令 $D'(X)=0$，则 $p=\frac{1}{2}$。

∴ 当 $p=\frac{1}{2}$时方差取得最大值，从而标准差最大，其最大值为$\sqrt{30}$。

3-15 设随机变量 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{\pi}}e^{-x^2-x-\frac{1}{4}},\quad -\infty < x < +\infty$$

求 $E(X)$，$E(X^2)$和 $D(2X-3)$。

【解】 $f(x)=\sqrt{2}\frac{1}{\sqrt{2\pi}}e^{\frac{-\left(x+\frac{1}{2}\right)^2}{1}}$，∴ $X \sim N\left(-\frac{1}{2},\frac{1}{2}\right)$

∴ $E(X)=-\frac{1}{2}$，$E(X^2)=D(X)+[E(X)]^2=\frac{3}{4}$

$D(2X-3)=E[(2X-3)^2]-[E(2X-3)]^2=-\frac{1}{2}$

3-16 设随机变量 X 的概率密度为

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

求 $E(X)$，$D(X)$。

【解】
$$E(X)=\int_0^1 x \times 2x\mathrm{d}x=\frac{2}{3}$$
$$D(X)=E(X^2)-[E(X)]^2=\int_0^1 x^2 \times 2x\mathrm{d}x-\left(\frac{2}{3}\right)^2=\frac{1}{18}$$

3-17 设随机变量 X 在区间$[-1,2]$上服从均匀分布，随机变量

$$Y=\begin{cases}1, & X>0\\ 0, & X=0\\ -1, & X<0\end{cases}$$

求方差 $D(Y)$。

【解】 依题意，可得 Y 的分布律表为

X	$[-1,0)$	0	$(0,2]$
Y	-1	0	1
P	$\frac{1}{3}$	0	$\frac{2}{3}$

$$E(Y)=(-1)\times\frac{1}{3}+0\times 0+1\times\frac{2}{3}=\frac{1}{3}$$
$$D(Y)=\left(-1-\frac{1}{3}\right)^2\times\frac{1}{3}+\left(1-\frac{1}{3}\right)^2\times\frac{2}{3}=\frac{8}{9}$$

3-18 假设有 10 个同种电器元件,其中有两个废品。装配仪器时,从这 10 个元件中任取一个,如果是废品,则扔掉重新任取一个;如果仍是废品,则扔掉再取一个。试求在取到正品之前,已取出的废品个数的概率分布、数学期望和方差。

【解】 由题意可知

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

$$E(X)=\frac{2}{10}\times\frac{8}{9}+\frac{1}{45}\times 2=\frac{2}{9}$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{2}{10}\times\frac{8}{9}+\frac{1}{45}\times 4-\left(\frac{2}{9}\right)^2=0.217\,27$$

3-19 已知离散型随机变量 X 的分布律为

X	-2	-1	0	2	1
p_k	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

令 $Y=X^2$,问 X 和 Y 是否不相关?X 和 Y 是否相互独立?

【解】

Y	4	1	0
P	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{1}{5}$

X \ Y	4	1	0
-2	$\frac{1}{5}$	0	0
-1	0	$\frac{1}{5}$	0
0	0	0	$\frac{1}{5}$
2	$\frac{1}{5}$	0	0
1	0	$\frac{1}{5}$	0

独立性:若 X 和 Y 相互独立,则

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$$

而 $P\{X=-2\}=\frac{1}{5}$,$P\{Y=4\}=\frac{2}{5}$,但

$$P\{X=-2,Y=4\}=\frac{1}{5}\neq P\{X=-2\}P\{Y=4\} \quad \therefore \text{不独立}$$

相关性：$E(X)=0, E(Y)=2, E(XY)=0$；$D(X)=2, D(Y)=\dfrac{24}{5}$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0$$

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=0 \quad \therefore X,Y\text{ 不相关}$$

达标实训Ⅱ

3-20 一台设备在一个工作日内出现故障的概率为 0.05，并且一旦出现故障就要全天停工进行维修。假设在一周 5 个工作日内无故障，可获利润 10 万元；发生一次故障仍可获利润 5 万元；发生两次故障获利润 0 元；发生三次和三次以上故障要亏损 2 万元。求一周 5 个工作日的平均利润。

【分析】 根据题意可知利润和故障次数的关系，由此可以得到概率分布律；再根据离散型数学期望的定义可以求出平均利润。

【解】 设获得利润的概率为 P，则 X 可能取值 $X=10,5,0,-2$（单位：万元）。

X	10	5	0	-2	-2	-2
P	$C_5^0\times0.95^5$	$C_5^1\times0.05\times0.95^4$	$C_5^2\times0.05^2\times0.95^3$	$C_5^3\times0.05^3\times0.95^2$	$C_5^4\times0.05^4\times0.95$	$C_5^5\times0.05^5$

$$E(X)=10\times C_5^0\times0.95^5+5\times C_5^1\times0.05\times0.95^4+0\times C_5^2\times0.05^2\times0.95^3$$
$$-2\times(C_5^3\times0.05^3\times0.95+C_5^4\times0.05^4\times0.95+C_5^5\times0.05^5)=8.7536$$

3-21 设 (X,Y) 的分布律为

Y \ X	1	2	3
-1	0.2	0.1	0
0	0.1	0	0.3
2	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 $E(Z)$。

【分析】 由题中联合分布，再根据离散型数学期望的定义可以得到随机变量的数学期望；由原有两个随机变量通过混合运算可得到新变量 Z，它的数学期望计算一般有两种方式，一种是根据原有两个变量的联合分布计算出新变量 Z 的每一种取值情况及其对应的概率，再利用离散型数学期望的定义计算即可，另一种是利用数学期望的运算性质，对所求的数学期望首先进行化简，在化简过程中对不知道的部分再分别进行求解。

【解】 (1) $E(X)=(0.2+0.1+0.1)\times1+(0.1+0+0.1)\times2+(0.1+0.3)\times3=2$

$E(Y)=(0.2+0.1+0)\times(-1)+(0.1+0+0.3)\times0+(0.1+0.1+0.1)\times2$
$=0.3$

(2) $E(Z)=1\times(-1)\times P\{X=1,Y=-1\}+\frac{1}{2}\times(-1)\times P\{X=2,Y=-1\}+\frac{1}{3}$

$\times(-1)\times P\{X=3,Y=-1\}+1\times0\times P\{X=1,Y=0\}+\frac{1}{2}\times0$

$\times P\{X=2,Y=0\}+\frac{1}{3}\times0\times P\{X=3,Y=0\}+1\times2$

$\times P\{X=1,Y=2\}+\frac{1}{2}\times2\times P\{X=2,Y=2\}+\frac{1}{3}\times2$

$\times P\{X=3,Y=2\}=-0.2+(-0.05)+0.2+0.1+\frac{2}{3}\times0.1$

$=\frac{0.2}{3}+0.05=\frac{7}{60}$

(3) $Z=(X-Y)^2=X^2+Y^2-2XY$

$E(X^2)=0.4\times1^2+0.2\times4+0.4\times9=4.8$

$E(Y^2)=0.3\times(-1)^2+0.4\times0^2+0.3\times2^2=1.5$

$E(2XY)=2E(XY)=2[1\times(-1)\times0.2+2\times(-1)\times0.1+3\times(-1)\times0+1\times0$
$\times0.1+2\times0\times0+3\times0\times0.3+1\times2\times0.1+2\times2\times0.1+3\times2\times0.1]$
$=2\times0.8=1.6$

$E(Z)=E[(X-Y)^2]=E(X^2)+E(Y^2)-E(2XY)=4.8+1.5-1.6=4.7$

3-22 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}e^{-x}, & x>0\\ 0, & x\leqslant0\end{cases}$$

求(1)$Y=2X$;(2)$Y=e^{-2X}$的数学期望。

【分析】 连续型随机变量数学期望,需要应用随机变量函数的数学期望的计算方法,积分即可。

【解】 (1) $E(2X)=\int_0^{+\infty}e^{-x}2x\mathrm{d}x=2(-xe^{-x}-e^{-x})\big|_0^{-\infty}=2$

(2) $E(e^{-2X})=\int_0^{+\infty}e^{-x}\cdot e^{-2x}\mathrm{d}x=\int_0^{+\infty}e^{-3x}\mathrm{d}x=-\frac{1}{3}e^{-3x}\big|_0^{+\infty}=\frac{1}{3}$

3-23 将 n 只球($1\sim n$ 号)随机放入 n 个盒子($1\sim n$ 号)中,且一个盒子只装一只球。若球号和盒子编号相同,则称为一个配对。记 X 为总的配对数,求 $E(X)$。

【分析】 随机变量 X 的所有可能取值为 $0,1,2,\cdots,n$,本题如果利用给出分布律,直接求期望的方法计算是非常复杂的,应利用期望的性质,将事件分成若干个小事件进行计算。

【解】 记 $X_i=\begin{cases}1, & \text{第 } i \text{ 个球恰好进 } i \text{ 盒}\\ 0, & \text{第 } i \text{ 个球未进 } i \text{ 盒}\end{cases}$, $i=1,2,\cdots,n$。

则 $X_1,X_2,\cdots,X_n$ 是同分布的,但不独立,其共同分布为

$$P(X_i=1)=\frac{1}{n},\quad P(X_i=0)=1-\frac{1}{n},\quad i=1,2,\cdots,n$$

由此得

$$E(X_i)=\frac{1}{n},\quad i=1,2,\cdots,n$$

又因为 $X=X_1+X_2+\cdots+X_n$，所以

$$E(X)=E(X_1)+E(X_2)+\cdots+E(X_n)=n\cdot\frac{1}{n}=1$$

3-24 某保险公司制订赔偿方案：如果在一年内一个顾客的投保事件 A 发生，该公司就赔偿顾客 a 元。若已知一年内事件 A 发生的概率为 p，为使保险公司收益的期望值等于 a 的 5%，则该公司应该要求顾客缴纳多少保险费？

【分析】 将顾客缴纳的保险费设为参数，根据收益期望等于 a 的 5%列方程求解出待定参数。

【解】 设应缴纳保险金 b 元，

X	b	$-a$
P	$1-p$	p

$$E(X)=(1-p)b-pa=5\%a$$

$$b=\frac{a(p+5\%)}{1-p}$$

3-25 设随机变量 X 服从参数为 2 的指数分布，求 $E(2X-\mathrm{e}^{-2X})$。

【分析】 由题意可知随机变量的密度函数，应用随机变量函数的数学期望的计算方法，积分可得。

【解】 $E(2X-\mathrm{e}^{-2X})=2E(X)-E(\mathrm{e}^{-2X})$

随机变量 X 服从参数为 2 的指数分布，$E(X)=\frac{1}{2}$。

$$E(\mathrm{e}^{-2X})=\int_0^{+\infty}2\mathrm{e}^{-2x}\mathrm{e}^{-2x}\mathrm{d}x=-\frac{1}{2}\mathrm{e}^{-4x}\Big|_0^{+\infty}=\frac{1}{2}$$

$\therefore$ $E(2X-\mathrm{e}^{-2X})=1-\frac{1}{2}=\frac{1}{2}$

3-26 一工厂生产的某种设备的寿命 X(以年计)服从参数为 $\lambda=\frac{1}{4}$ 的指数分布，概率密度为

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

工厂规定出售的设备如果在售出一年内损坏就予以调换，若工厂售出一台设备盈利 100 元，调换一台设备厂方花费 300 元，试求厂方出售一台设备净盈利的数学期望。

【分析】 根据设备寿命情况首先得到寿命在一年内和一年以上的概率，再根据数学期望的定义去求盈利的数学期望。

【解】 设厂方出售一台设备净盈利为 Y，Y 的可能取值为 100，-300。

$$P\{Y=100\}=P\{X\geqslant 1\}=\int_1^{+\infty}\frac{1}{4}\mathrm{e}^{-\frac{x}{4}}\mathrm{d}x=\mathrm{e}^{-\frac{1}{4}}$$

$$P\{Y=300\}=P\{X<1\}=\int_0^1 \frac{1}{4}\mathrm{e}^{-\frac{x}{4}}\mathrm{d}x=1-\mathrm{e}^{\frac{1}{4}}$$

所以,厂方出售一台设备的净盈利的数学期望为

$$E(Y)=100\mathrm{e}^{-\frac{1}{4}}-300(1-\mathrm{e}^{-\frac{1}{4}})=400\mathrm{e}^{-\frac{1}{4}}-300$$

3-27 假设某种热水器首次发生故障的时间 X(单位:h)服从参数为 $\lambda=\frac{1}{500}$ 的指数分布,求(1)该热水器在 100h 内需要维修的概率是多少?(2)该热水器平均能正常使用多长时间?

【分析】 根据指数分布的密度函数,应用连续型随机变量函数的数学期望的方法,在指定区域内积分即可。

【解】 (1) 由题意解得

$$f(x)=\begin{cases}\frac{1}{500}\mathrm{e}^{-\frac{1}{500}x}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

$$\therefore\ P\{X\leqslant 100\}=\int_0^{100}\frac{1}{500}\mathrm{e}^{-\frac{1}{500}x}\mathrm{d}x=1-\mathrm{e}^{-\frac{1}{5}}$$

(2) $E(X)=\frac{1}{\lambda}=500$

∴ 该热水器在 100h 内需要维修的概率为 $1-\mathrm{e}^{-\frac{1}{5}}$;该热水器平均正常使用 500h。

3-28 假设自动加工的某种零件的内径 X(单位:mm)服从正态分布 $N(\mu,1)$,内径小于 10mm 或大于 12mm 为不合格品,其余为合格品,销售每件合格品获利,销售每件不合格品亏损,已知销售利润 T(单位:元)与销售零件的内径之间有以下关系。

$$T=\begin{cases}-1, & x<10\\ 20, & 10\leqslant x\leqslant 12\\ -25, & x>12\end{cases}$$

问平均内径 μ 取何值时,销售一个零件平均利润最大?

【分析】 随机变量属于连续型,因此利用正态分布的密度函数积分求出含有参数 μ 的利润的数学期望,接下来讨论此数学期望函数的单调性和最值,可求得使得销售一个零件平均利润最大的 μ 取值。

【解】 $X\sim N(\mu,1)$,则 $X-\mu\sim N(0,1)$。

$$\therefore\ T=\begin{cases}-1, & X-\mu<10-\mu\\ 20, & 10-\mu\leqslant X-\mu\leqslant 12-\mu\\ -25, & X-\mu>12-\mu\end{cases}$$

$$E(T)=(-1)\int_{-\infty}^{10-\mu}\varphi(x)\mathrm{d}x+20\int_{10-\mu}^{12-\mu}\varphi(x)\mathrm{d}x-25\int_{12-\mu}^{+\infty}\varphi(x)\mathrm{d}x$$

$E(T)$ 为 μ 的函数,设该函数为 $F(\mu)$,则 $F'(\mu)=214(10-\mu)-45\varphi(12-\mu)$。

令 $F'(\mu)=0\Rightarrow\mu=10.62$

当 $\mu>10.62$ 时,$F'(\mu)<0$,$F(\mu)$ 单调递减;

当 $\mu<10.62$ 时,$F'(\mu)>0$,$F(\mu)$ 单调递增,所以当 $\mu=10.62$ 时,$E(T)$ 最大。

3-29　设随机变量 X 服从参数为 λ 的泊松分布，且已知 $E[(X-1)(X-2)]=1$，计算 $P\{X\neq 0\}$。

【分析】　首先由已知的条件利用数学期望的性质求出泊松分布的 λ，再去计算所求概率。

【解】　$E[(X-1)(X-2)]=1$，即 $E(X^2)-3E(X)+2=1$。
由已知可得 $E(X)=\lambda, D(X)=\lambda$
$\therefore\ \lambda^2+\lambda-3\lambda+2=1\Rightarrow\lambda=1$

$$p\{X\neq 0\}=1-p\{X=0\}=1-\frac{1^0}{0!}=1-\frac{1}{\mathrm{e}}=0.632$$

3-30　一水果商每天卖出去的水果数量(单位：kg)服从参数为 λ 的泊松分布。如果每卖出 1kg 可得报酬 a 元，卖不掉则损失 b 元，若每天水果商批发进水果 $n$kg，求其期望所得和最佳的买进数量 n。

【分析】　待定参数 n，首先根据题意利用数学期望的定义，可解出期望所得。

【解】　由题意知利润为

$$a\cdot\frac{\lambda k}{k!}\mathrm{e}^{-\lambda}-\left(n-\frac{\lambda k}{k!}\mathrm{e}^{-\lambda}\right)b$$

$$E(X)=(a-b)\lambda$$

3-31　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，而积压一件产品会导致 n 元的损失。再者，他们预测销售量 Y(单位：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)？

【分析】　首先根据题意可以得到利润函数，对利润函数求数学期望，会得到关于 x 的函数表达式，通过对此表达式求导，解除最值对应的 x 即可。

【解】　设生产 x 件，则获利 Q 是 x 的函数。

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

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

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

$$x=-\theta\ln\left(\frac{n}{m+n}\right)$$

又因为 $\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}E(Q)=\dfrac{-(m+n)}{\theta}\mathrm{e}^{-\frac{x}{\theta}}<0$

因此，当 $x=-\theta\ln\left(\frac{n}{m+n}\right)$ 时，$E(Q)$ 取最大值。

3-32 一台设备由三个部件构成，在设备运行过程中，各部件需要调整的概率分别为 0.1，0.2，0.3。假设各部件的状态相互独立，以 X 表示同时需要调整的部件个数，求 X 的方差。

【分析】 依据题意可求随机变量 X 的分布律，根据分布律可求得 X 和 X^2 的数学期望，进而得到 X 的方差。$P\{x=0\}=0.9\times0.8\times0.7=0.504$

【解】 由题意可知

$$P\{X=1\}=0.1\times0.8\times0.7+0.9\times0.2\times0.7+0.9\times0.8\times0.3$$
$$P\{X=2\}=0.1\times0.2\times0.7+0.1\times0.8\times0.3+0.9\times0.2\times0.3$$
$$P\{X=3\}=0.1\times0.2\times0.3$$

即有

X	0	1	2	3
P	0.504	0.398	0.092	0.006

因此，

$$E(X)=1\times0.398+2\times0.092+3\times0.006=0.6$$
$$E(X^2)=1\times0.398+4\times0.092+9\times0.006=0.82$$
$$D(X)=E(X^2)-[E(X)]^2=0.46$$

3-33 掷 n 颗骰子，骰子每一面的出现都是等可能的，求出现点数之和的期望和方差。

【分析】 根据离散型随机变量数学期望的定义求点数之和 X 与 X^2 的期望，进而得到方差。

【解】 设 X_k 表示第 k 次掷骰子的数字，则 $X=X_1+X_2+\cdots+X_k$，且 $X_1,X_2,\cdots,X_k$ 独立，则

$$P\{X_i=j\}=\frac{1}{6},\quad j=1,2,3\cdots,6;i=1,2,3,\cdots,n$$

$$P\{X_i=j\}=\frac{1}{6}$$

$$E(X_i)=\sum_{j=1}^{6}j\times\frac{1}{6}=\frac{1}{6}\times\frac{(1+6)\times6}{2}=\frac{7}{2}$$

$$E(X^2)=\sum_{j=1}^{6}j^2\times\frac{1}{6}=(1^2+2^2+\cdots+6^2)\times\frac{1}{6}=\frac{91}{6}$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{161}{12}$$

因此，

$$P(X_j)=\frac{1}{6}n$$

$$E(X_i)=\frac{7}{2}n$$

$$D(X_i)=\frac{161}{12}n$$

3-34　设二维随机变量(X,Y)的概率密度为

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

验证：X 和 Y 是相关的，但 X 和 Y 不是相互独立的。

【分析】 由概率密度 $f(x,y)=\begin{cases}\dfrac{1}{\pi}, & x^2+y^2\leqslant 1\\ 0, & 其他\end{cases}$可知，二维随机变量$(X,Y)$服从单位圆上的均匀分布。首先求出 X 与 Y 的边缘概率密度，验证两者的乘积是否等于联合概率密度，并判断独立性。然后，求 X 与 Y 的协方差，进而求相关系数。

【解】 独立性：X 的边缘概率密度为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\displaystyle\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y, & -1\leqslant x\leqslant 1\\ 0, & 其他\end{cases}$$

$$=\begin{cases}\dfrac{2}{\pi}\sqrt{1-x^2}, & -1\leqslant x\leqslant 1\\ 0, & 其他\end{cases}$$

Y 的边缘概率密度为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\displaystyle\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}\mathrm{d}x, & -1\leqslant y\leqslant 1\\ 0, & 其他\end{cases}$$

$$=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2}, & -1\leqslant y\leqslant 1\\ 0, & 其他\end{cases}$$

显然，$f(x,y)\neq f_X(x)f_Y(y)$，故 X 和 Y 不是相互独立的。

相关性：$\mathrm{Cov}(X,Y)=\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}\mathrm{d}x\mathrm{d}y$

$$=\frac{2}{\pi}\int_{-1}^{1}\sqrt{1-y^2}\mathrm{d}y=1\neq 0$$

故相关系数 $\rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}\neq 0$，故 X 和 Y 是相关的。

3-35　设二维随机变量(X,Y)的联合概率密度为

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

求(X,Y)的相关系数 ρ_{XY}。

【分析】 在联合概率密度已知的情况下很容易求得变量的期望、方差和协方差，再根据相关系数的公式求解即可。

【解】 $E(X)=\int_0^2\int_0^2 y\cdot\frac{1}{8}(x+y)\mathrm{d}x\mathrm{d}y=\frac{7}{6}$

$$E(Y)=\int_0^2\int_0^2 y\cdot\frac{1}{8}(x+y)\mathrm{d}x\mathrm{d}y=\frac{7}{6}$$

$$\begin{aligned}\mathrm{Cov}(X,Y)&=\int_0^2\int_0^2[x-E(X)][y-E(Y)]f(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_0^2\int_0^2\left(x-\frac{7}{6}\right)\left(y-\frac{7}{6}\right)\times\frac{1}{8}(x+y)\mathrm{d}x\mathrm{d}y=-\frac{1}{36}\end{aligned}$$

$$E(XY)=\mathrm{Cov}(X,Y)+E(X)E(Y)=-\frac{1}{36}+\frac{49}{36}=\frac{4}{3}$$

$$D(X)=\int_0^2\left(x-\frac{7}{6}\right)^2 f(x)\mathrm{d}x=\int_0^2\left(x-\frac{7}{6}\right)^2\left(\frac{1+x}{4}\right)\mathrm{d}x=\frac{11}{36}$$

$$D(Y)=\int_0^2\left(y-\frac{7}{6}\right)^2 f(y)\mathrm{d}y=\int_0^2\left(y-\frac{7}{6}\right)^2\left(\frac{1+y}{4}\right)\mathrm{d}y=\frac{11}{36}$$

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=-\frac{1}{11}$$

3-36 设在每次试验中，事件 A 发生的概率为 0.5，利用切比雪夫不等式估计在 1000 次独立试验中，事件 A 发生的次数为 400～600 的概率。

【分析】 考查切比雪夫不等式的知识点。

【解】 $E(X)=np=500$

$$P\{|X-500|<100\}\geqslant 1-\frac{D(X)}{\varepsilon^2}$$

$$D(X)=np(1-p)=250$$

$$P\{|X-500|<100\}\geqslant 1-\frac{250}{10\ 000}$$

$\therefore\ P\{|X-500|<100\}\geqslant\frac{39}{40}$

3-37 一个零件的重量是一个随机变量，其期望值是 1kg，标准差是 0.1kg，求一箱（100 个）相同零件的重量超过 102kg 的概率。

【分析】 根据中心定理大量独立事件之和考虑服从正态分布，进而求解。

【解】 假设 X_i 表示第 i 个零件的变量 $i=1,\cdots,100$，则 $X_i,\cdots,X_{100}$ 相互独立同分布，$E(X_i)=1,D(X_i)=0.1^2$，记一箱相同零件的重量为 S_{100}，

$$S_{100}=\sum_{i=1}^{100}x_i\quad E_{S_{100}}=100,\quad D_{S_{100}}=1$$

由中心定理可知 S_{100} 近似服从正态分布 $N(100,1^2)$，
即

$$\begin{aligned}P\{S_{100}>102\}&=1-P\{S_{100}\leqslant 102\}=1-P\left\{\frac{S_{100}-100}{1}\leqslant 2\right\}\\&=1-\Phi(2)=0.022\ 75\end{aligned}$$

所以一箱相同零件的重量超过 102kg 的概率为 0.022 75。

3-38　甲乙两个剧院在竞争 1000 名观众。假设每个观众完全随意地选择一个剧院，且观众选择剧院是彼此独立的，问每个剧院应设多少个座位，才能保证因缺少座位而使观众离去的概率小于 1%？

【分析】　考查中心极限定理。

【解】　设 X 为进入某个剧院的人数，由题意得

$$X \sim B\left(1000, \frac{1}{2}\right)$$

由棣莫弗-拉普拉斯中心极限定理有 X 近似服从 $N(500,250)$ 的正态分布。

设 x 为该剧院所设座位数，则

$$P\{X \leqslant x\} = 0.99 \quad 即 \quad P\left\{\frac{X-500}{\sqrt{250}} \leqslant \frac{x-500}{\sqrt{250}}\right\} = 0.99$$

即　$\Phi\left(\frac{x-500}{\sqrt{250}}\right)=0.99 \Rightarrow \frac{x-500}{\sqrt{250}}=2.33 \Rightarrow x=537$

3-39　卡车装运水泥，设每袋水泥的重量 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 2000 的概率不大于 0.05。

【分析】　考查中心极限定理。

【解】　假设装 i 袋水泥使总质量超过 2000 的概率小于 0.5，则 $x_i \cdots x_{100}$ 相互独立同分布，记 i 袋水泥的质量为 S_i，则 $E(S_i)=50i, D(S_i)=2.5^2 i$。

由中心定理可知，S_i 近似服从正态分布 $N(50i, 2.5^2 i)$，则

$$P\{S_i > 2000\} = 1 - P\{S_i \leqslant 2000\}$$

$\because P\{S_i>2000\} \leqslant 0.05$，则 $P\{S_i \leqslant 2000\} > 0.05$

即 $P\{S_i \leqslant 2000\} = P\left\{\frac{S_i-50i}{2.5\sqrt{i}} \leqslant \frac{2000-50i}{2.5\sqrt{i}}\right\} = \Phi\left(\frac{2000-50i}{2.5\sqrt{i}}\right)$

若 $\Phi\left(\frac{2000-50i}{2.5\sqrt{i}}\right)=0.05$，则 $i=40$ 袋。

若 $\Phi\left(\frac{2000-50i}{2.5\sqrt{i}}\right)>0.05$，则 $\frac{2000-50i}{2.5\sqrt{i}}>0$。

$\therefore i<40$，i 取 39 袋。

$\therefore$ 最多装 39 袋水泥使总质量超过 2000 的概率不大于 0.05。

达标自测题

1. 如果 X,Y 相互独立，则（　　）。

A. $D(XY)=D(X)D(Y)$　　B. $D(X+3Y)=D(X)+3D(Y)$

C. $D(3X-2Y)=9D(X)-4D(Y)$　　D. $D(3X-2Y)=9D(X)+4D(Y)$

2. 设随机向量 (X,Y) 的联合分布密度为 $f(x,y)=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+(y-1)^2}{2}}, -\infty<x,y<+\infty$，

则(　　)。

A. (X,Y) 服从指数分布　　B. X 与 Y 不独立

C. X 与 Y 相互独立　　D. $\mathrm{Cov}(X,Y)\neq 0$

3. 若 $D(X)=4$,$D(Y)=1$,$\rho_{XY}=0.6$,则 $D(X+Y)=$(　　)。

A. 7.4　　B. 40　　C. 5.6　　D. 7.6

4. 设 $X\sim N(1,2)$,$Y\sim N(1,3)$,且 X 与 Y 相互独立,则 $3X-2Y\sim$(　　)。

A. $N(1,30)$　　B. $N(30,1)$　　C. $N(1,20)$　　D. $N(0,1)$

5. 设随机变量 X 的概率密度 $f(x)=\sqrt{\dfrac{3}{\pi}}\mathrm{e}^{-3x^2-12x-12}$,则 X 的方差 $D(X)=$(　　)。

A. $\dfrac{1}{6}$　　B. $\dfrac{1}{2}$　　C. $\dfrac{1}{3}$　　D. $\dfrac{1}{9}$

6. 设随机变量 X 与 Y 相互独立,$X\sim N(1,4)$,Y 服从参数为 1/3 的泊松分布,记 $Z=\dfrac{X}{2}-\dfrac{Y}{4}$,则协方差 $\mathrm{Cov}(X,Z)=$(　　)。

A. 1　　B. 2　　C. $\dfrac{1}{2}$　　D. 0

7. 设随机变量 X 服从二项分布 $X\sim B\left(4,\dfrac{1}{4}\right)$,则 X 的期望与方差 $\{E(X),D(X)\}=$(　　)。

A. $\left\{1,\dfrac{3}{8}\right\}$　　B. $\left\{\dfrac{1}{2},\dfrac{3}{8}\right\}$　　C. $\left\{\dfrac{1}{2},\dfrac{1}{4}\right\}$　　D. $\left\{1,\dfrac{3}{4}\right\}$

8. 设随机变量 X 服从区间 (a,b) 上的均匀分布 $(a<b)$,并且 $E(X)=1$,$D(X)=\dfrac{1}{3}$,则 a,b 的值为(　　)。

A. $a=0,b=4$　　B. $a=0,b=2$　　C. $a=1,b=3$　　D. $a=1,b=4$

9. 设随机变量 X 在区间 $[-3,3]$ 服从均匀分布,则 $E(X)=$(　　)。

A. 0　　B. 1　　C. 2　　D. 3

10. 100 个灯泡的寿命 $X_1,X_2,X_3,\cdots,X_{99},X_{100}$ 独立同分布,且 $E(X_i)=b$ $(i=1,2,3,\cdots,100)$,则 100 个灯泡的平均寿命 $Y=\dfrac{1}{100}(X_1+X_2+X_3+\cdots+X_{100})$ 的期望 $E(Y)=$(　　)。

A. $100b$　　B. b　　C. $0.01b$　　D. $0.04b$

11. 设 $\xi\sim N(3,4)$,η 在区间 $[0,10]$ 上服从均匀分布,则下列式子错误的是(　　)。

A. $E(\xi+\eta)=8$　　B. $D(\xi+\eta)=8$

C. $E\left(\xi^2+\eta^2-\dfrac{1}{3}\right)=46$　　D. $E\left(\dfrac{\xi}{2}+\dfrac{\eta}{5}-\dfrac{5}{2}\right)=0$

12. 设随机变量序列 $X_1,X_2,\cdots,X_n$ 独立服从同一分布,且方差 $\sigma^2>0$,$Y=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$,则下列各式成立的是(　　)。

A. $\mathrm{Cov}(X_1,Y)=n\sigma^2$　　B. $\mathrm{Cov}(X_1,Y)=\sigma^2$

C. $D(X_1+Y)=\dfrac{(n+2)\sigma^2}{n}$　　D. $D(X_1-Y)=\dfrac{(n-1)\sigma^2}{n}$

13. 设 5 个随机变量 X_1,X_2,X_3,X_4,X_5 独立同分布，且 $E(X_i)=a,D(X_i)=b,i=1,2,3,4,5$，设 $Y=\dfrac{1}{5}(X_1+X_2+X_3+X_4+X_5)$，则 $D(Y)=$(　　)。

A. $5b$　　B. b　　C. $0.2b$　　D. $0.04b$

14. 若二维随机变量 (X,Y) 的概率密度为 $f(x,y)=\begin{cases}\dfrac{1}{8}(x+y), & 0\leqslant x\leqslant 2,0\leqslant y\leqslant 2\\ 0, & 其他\end{cases}$ 则 $E(XY)=$(　　)。

A. $\dfrac{5}{6}$　　B. $\dfrac{4}{3}$　　C. 1　　D. $\dfrac{6}{7}$

15. 若二维随机变量 (X,Y) 的概率密度为 $f(x,y)=\begin{cases}\dfrac{1}{8}(x+y), & 0\leqslant x\leqslant 2,0\leqslant y\leqslant 2\\ 0, & 其他\end{cases}$，则 $\mathrm{Cov}(X,Y)=$(　　)。

A. $\dfrac{7}{6}$　　B. $-\dfrac{1}{36}$　　C. $\dfrac{1}{6}$　　D. $\dfrac{1}{36}$

16. 对于任意两个随机变量 ξ 和 η，若 $E(\xi\eta)=E(\xi)E(\eta)$，则有(　　)。

A. $D(\xi\eta)=D(\xi)D(\eta)$　　B. $D(\xi+\eta)=D(\xi)+D(\eta)$

C. ξ 和 η 独立　　D. ξ 和 η 不独立

17. 设随机变量 X,Y 独立同分布，记 $\xi=X+Y,\eta=X-Y$，则随机变量 ξ 和 η 之间的关系必然是(　　)。

A. 不独立　　B. 独立

C. 相关系数等于 0　　D. 相关系数不为 0

18. 设 ξ 的分布律为：$P\{\xi=n\}=P\{\xi=-n\}=\dfrac{1}{2n(n+1)}$，($n$ 为正整数)，则 $E(\xi)=$(　　)。

A. 0　　B. 1　　C. 0.5　　D. 不存在

19. 设随机变量 ξ_n，服从二项分布 $B(n,p)$，其中 $0<p<1,n=1,2,\cdots$ 那么，对于任一实数 x 有 $\lim\limits_{n\to+\infty}P\left\{\dfrac{\xi_n-np}{\sqrt{np(1-p)}}<x\right\}$ 等于(　　)。

A. $\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$　　B. 0

C. $\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt$　　D. $\displaystyle\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$

20. 设随机变量 ξ 的数学期望 $E(\xi)=\mu$，方差 $D(\xi)=\sigma^2$，试利用切比雪夫不等式估计 $P\{|\xi-\mu|<4\sigma\}\geqslant$(　　)。

A. $\dfrac{8}{9}$　　B. $\dfrac{15}{16}$　　C. $\dfrac{9}{10}$　　D. $\dfrac{1}{10}$

3.4 拓展实训

随机变量的数字特征是描述随机变量特征的有效工具，特别是几种最重要的数字特征：数学期望，方差，相关系数等都有明确的概率意义的同时，又具有良好的性质，因此数字特征的概念在概率论与数理统计中占据着重要的地位，在历年考研试题中出现的频率也相当高。而大数定律和中心极限定理主要包括一个不等式，两个定理和三个定律，在历年研究生入学考试中出现的并不多，总体来看这部分在近11年的考研数学中约占到概率统计部分题量29%的比例，历年考研试题的考点大致可归纳为：

(1) 求随机变量的数字特征；可归结为求随机变量及其函数的数学期望问题。

(2) 求两个随机变量的协方差和相关系数。

(3) 切比雪夫不等式。

(4) 棣莫弗-拉普拉斯定理；列维-林德伯格定理(独立同分布的中心极限定理)。

(5) 切比雪夫大数定律；伯努利大数定律；辛钦大数定律。

拓展实训 Ⅰ

3-40 【2003数三，填空题(5)，4分】设随机变量 X 和 Y 的相关系数为0.9，若 $Z=X-0.4$，则 Y 与 Z 的相关系数为__0.9__。

【分析】 利用相关系数的计算公式即可。

【详解】 因为

$$\begin{aligned}\mathrm{Cov}(Y,Z)&=\mathrm{Cov}(Y,X-0.4)=E[Y(X-0.4)]-E(Y)E(X-0.4)\\&=E(XY)-0.4E(Y)-E(Y)E(X)+0.4E(Y)\\&=E(XY)-E(X)E(Y)=\mathrm{Cov}(X,Y)\end{aligned}$$

且 $D(Z)=D(X)$

于是有 $\rho_{YZ}=\dfrac{\mathrm{Cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\rho_{XY}=0.9$

3-41 【2003数三，填空题(6)，4分】设总体 X 服从参数为2的指数分布，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本，则当 $n\to\infty$ 时，$Y_n=\dfrac{1}{n}\sum_{i=1}^{n}X_i^2$ 依概率收敛于__$\dfrac{1}{2}$__。

【分析】 本题考查大数定律：一组相互独立且具有有限期望与方差的随机变量 $X_1,X_2,\cdots,X_n$，当方差一致有界时，其算术平均值依概率收敛于其数学期望的算术平均值。

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{p}\frac{1}{n}\sum_{i=1}^{n}EX_i\ (n\to\infty)$$

【详解】 这里 $X_1^2,X_2^2,\cdots,X_n^2$ 满足大数定律的条件，且 $E(X_i^2)=D(X_i)+[E(X_i)]^2=$

$\frac{1}{4}+\left(\frac{1}{2}\right)^2=\frac{1}{2}$，因此根据大数定律有 $Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 依概率收敛于 $\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)=\frac{1}{2}$。

3-42 【2003 数四，填空题(6)，4 分】设随机变量 X 和 Y 的相关系数为 0.5，$E(X)=E(Y)=0$，$E(X^2)=E(Y^2)=2$，则 $E(X+Y)^2=$___6___。

【分析】 利用期望与相关系数的公式进行计算即可。

【详解】 因为

$$\begin{aligned}E(X+Y)^2&=E(X^2)+2E(XY)+E(Y^2)\\&=4+2[\mathrm{Cov}(X,Y)+E(X)E(Y)]\\&=4+2\rho_{XY}\sqrt{D(X)}\cdot\sqrt{D(Y)}=4+2\times0.5\times2=6\end{aligned}$$

【评注】 本题的核心是逆向思维，利用公式 $E(XY)=\mathrm{Cov}(X,Y)+E(X)E(Y)$。

3-43 【2003 数四，选择题(6)，4 分】设随机变量 X 和 Y 都服从正态分布，且它们不相关，则(　　)。

A. X 与 Y 一定相互独立　　B. (X,Y)服从二维正态分布

C. X 与 Y 未必相互独立　　D. $X+Y$ 服从一维正态分布

【分析】 本题考查正态分布的性质以及二维正态分布与一维正态分布之间的关系。只有(X,Y) 服从二维正态分布时，不相关与独立才是等价的。

【详解】 只有当(X,Y)服从二维正态分布时，X 与 Y 不相关$\Leftrightarrow X$ 与 Y 独立，本题仅仅已知 X 和 Y 服从正态分布，因此，由它们不相关推不出 X 与 Y 一定独立，排除 A；若 X 和 Y 都服从正态分布且相互独立，则(X,Y)服从二维正态分布，但题设并不知道 X,Y 是否独立，可排除 B；同样要求 X 与 Y 相互独立时，才能推出 $X+Y$ 服从一维正态分布，可排除 D。故正确选项为 C。

【评注】 ① 若 X 与 Y 均服从正态分布且相互独立，则(X,Y)服从二维正态分布。

② 若 X 与 Y 均服从正态分布且相互独立，则 $aX+bY$ 服从一维正态分布。

③ 若(X,Y)服从二维正态分布，则 X 与 Y 相互独立$\Leftrightarrow X$ 与 Y 不相关。

3-44 【2004 数一，(6)；2004 数三，(5)；2004 数四，(6)，4 分】设随机变量 X 服从参数为 λ 的指数分布，则 $P\{X>\sqrt{D(X)}\}=$___$\frac{1}{\mathrm{e}}$___。

【分析】 已知连续型随机变量 X 的分布，求其满足一定条件的概率，转化为定积分计算即可。

【详解】 由题设，知 $D(X)=\frac{1}{\lambda^2}$，于是

$$P\{X>\sqrt{DX}\}=P\left\{X>\frac{1}{\lambda}\right\}=\int_{\frac{1}{\lambda}}^{+\infty}\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x=-\mathrm{e}^{-\lambda x}\Big|_{\frac{1}{\lambda}}^{+\infty}=\frac{1}{\mathrm{e}}$$

【评注】 本题属于基本题型，应记住常见指数分布等的期望与方差的数字特征，而不应在解题时再去推算。

3-45 【2004 数四，(14)，4 分】设随机变量 $X_1,X_2,\cdots,X_n(n>1)$ 独立同分布，且方差 $\sigma^2>0$。令随机变量 $Y=\frac{1}{n}\sum_{i=1}^{n}X_i$，则(　　)。

A. $D(X_1+Y)=\frac{n+2}{n}\sigma^2$　　　　B. $D(X_1-Y)=\frac{n+2}{n}\sigma^2$

C. $\mathrm{Cov}(X_1,Y)=\frac{\sigma^2}{n}$　　　　D. $\mathrm{Cov}(X_1,Y)=\sigma^2$

【分析】 利用协方差的性质立即得正确答案。

【详解】 由于随机变量 $X_1,X_2,\cdots,X_n(n>1)$独立同分布，于是可得

$$\begin{aligned}\mathrm{Cov}(X_1,Y)&=\mathrm{Cov}\left(X_1,\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\mathrm{Cov}\left(X_1,\sum_{i=1}^{n}X_i\right)\\&=\frac{1}{n}\sum_{i=1}^{n}\mathrm{Cov}(X_1,X_i)=\frac{1}{n}\mathrm{Cov}(X_1,X_1)\\&=\frac{1}{n}D(X_1)=\frac{1}{n}\sigma^2\end{aligned}$$

故正确答案为 C。

【评注】 本题是对协方差性质的考查，属于基本题。

3-46 【2005 数四，(14)，4 分】设 $X_1,X_2,\cdots,X_n,\cdots$为独立同分布的随机变量列，且均服从参数为 $\lambda(\lambda>1)$的指数分布，记 $\Phi(x)$为标准正态分布函数，则(　　)。

A. $\lim\limits_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-n\lambda}{\lambda\sqrt{n}}\leqslant x\right\}=\Phi(x)$　　　　B. $\lim\limits_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-n\lambda}{\sqrt{n\lambda}}\leqslant x\right\}=\Phi(x)$

C. $\lim\limits_{n\to\infty}P\left\{\frac{\lambda\sum_{i=1}^{n}X_i-n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$　　　　D. $\lim\limits_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-\lambda}{\sqrt{n}\lambda}\leqslant x\right\}=\Phi(x)$

【分析】 只需求出 $\sum_{i=1}^{n}X_i$ 的期望与方差，再根据中心极限定理将其标准化即可。

【详解】 由题设，$E(X_i)=\frac{1}{\lambda},D(X_i)=\frac{1}{\lambda^2},i=1,2,\cdots,n,\cdots$于是

$$E\left(\sum_{i=1}^{n}X_i\right)=\frac{n}{\lambda},\quad D\left(\sum_{i=1}^{n}X_i\right)=\frac{n}{\lambda^2}$$

根据中心极限定理，知$\dfrac{\sum_{i=1}^{n}X_i-\frac{n}{\lambda}}{\sqrt{\frac{n}{\lambda^2}}}=\dfrac{\lambda\sum_{i=1}^{n}X_i-n}{\sqrt{n}}$其极限分布服从标准正态分布，故应选 C。

3-47 【2008 数一、数三、数四，(8)，4 分】设随机变量 $X\sim N(0,1),Y\sim N(1,4)$且相关系数 $\rho_{XY}=1$，则(　　)。

A. $P\{Y=-2X-1\}=1$　　　　B. $P\{Y=2X-1\}=1$

C. $P\{Y=-2X+1\}=1$　　　　D. $P\{Y=2X+1\}=1$

【分析】 用排除法。

设 $Y=aX+b$，由 $\rho_{XY}=1$，知道 X,Y 正相关，得 $a>0$，排除 A、C。

由 $X\sim N(0,1),Y\sim N(1,4)$，得

$$E(X)=0,\quad E(Y)=1,\quad E(Y)=E(aX+b)=aE(X)+b$$

$1=a\times0+b$，$b=1$，排除 B。

故选择 D。

3-48 【2008 数一、数四，(14)，4 分】设随机变量 X 服从参数为 1 的泊松分布，则 $P\{X=E(X^2)\}=$________。

【分析】 因为 $D(X)=E(X^2)-(EX)^2$，所以 $E(X^2)=2$，X 服从参数为 1 的泊松分布，所以 $P\{X=2\}=\frac{1}{2}e^{-1}$。

3-49 【2009 数一，(7)，4 分】设随机变量 X 的分布函数为 $F(x)=0.3\Phi(x)+0.7\Phi\left(\frac{x-1}{2}\right)$，其中 $\Phi(x)$为标准正态分布函数，则 $E(X)=($ $)$。

A. 0 B. 0.3 C. 0.7 D. 1

【分析】 因为 $F(x)=0.3\Phi(x)+0.7\Phi\left(\frac{x-1}{2}\right)$

所以 $F'(x)=0.3\varphi(x)+\frac{0.7}{2}\varphi\left(\frac{x-1}{2}\right)=0.3\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+0.7\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2\times2^2}}$

由于$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$是 $N(0,1)$的密度函数，故其期望值为 0，

$\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2\times2^2}}$ 是 $N(1,2^2)$ 的密度函数，其期望值为 1，所以

$E(X)=\int_{-\infty}^{+\infty}xF'(x)dx=0.3\times0+0.7\times1=0.7$，故选 C。

【点评】 这是一个已知分布函数求期望的问题，属于概率论的基本题型。需要知道正态分布的基本性质，这类问题在辅导讲义中有许多类似的题目。

3-50 【2010 数一，(14)，4 分】设随机变量 X 概率分布为 $p\{X=k\}=\frac{C}{k!}$，$k=0,1,2,\cdots$，则$E(X^2)=$ ________。

【详解】 由概率密度的性质$\sum_{k=0}^{\infty}P\{X=k\}=1$，有$\sum_{k=0}^{\infty}\frac{C}{k!}=1\Rightarrow C=e^{-1}$ 。

即 $P\{X=k\}=\frac{e^{-1}}{k!}$，$k=0,1,2,\cdots$为参数为 1 的泊松分布，则有

$$E(X)=1,\quad D(X)=1\Rightarrow E(X^2)=D(X)+[E(X)]^2=2$$

3-51 【2011 年数一，(8)，4 分】设随机变量 X 与 Y 相互独立，且 $E(X)$与 $E(Y)$存在，记 $U=\max\{X,Y\}$，$V=\min\{X,Y\}$，则 $E(UV)=($ $)$。

A. $E(U)E(V)$ B. $E(X)E(Y)$ C. $E(U)E(Y)$ D. $E(X)E(V)$

【分析】 本题考查随机变量数字特征的运算性质。计算时需要先对随机变量 UV 进行处理，有一定的灵活性。

【详解】 由于 $UV=\max\{X,Y\}\min\{X,Y\}=XY$

可知 $E(UV)=E[\max(X,Y)\min(X,Y)]=E(XY)=E(X)E(Y)$

故应选 B。

3-52 【2011 年数一、数三,(14),4 分】设二维随机变量(X,Y)服从$N(\mu,\mu;\sigma^2,\sigma^2;0)$,则$E(XY^2)=$________。

【分析】 本题考查二维正态分布的性质。

【详解】 由于$\rho=0$,由二维正态分布的性质可知随机变量X,Y独立。因此$E(XY^2)=E(X)E(Y^2)$。

由于(X,Y)服从$N(\mu,\mu;\sigma^2,\sigma^2;0)$,所以$X\sim N(\mu,\sigma^2)$,$Y\sim N(\mu,\sigma^2)$。

可知$E(X)=\mu$,$E(Y^2)=D(Y)+(EY)^2=\mu^2+\sigma^2$,又因为$\rho=0$,所以$X,Y$独立,则$E(XY^2)=E(X)E(Y^2)=\mu(\mu^2+\sigma^2)=\mu^3+\mu\sigma^2$。

3-53 【2012 年数一,(7),4 分】设随机变量X与Y相互独立,且分别服从参数为 1 与参数为 4 的指数分布,则$P\{X<Y\}=(\quad)$。

A. $\frac{1}{5}$　　B. $\frac{1}{3}$　　C. $\frac{2}{5}$　　D. $\frac{4}{5}$

【分析】 (X,Y)的联合概率密度为$f(x,y)=\begin{cases}e^{-x-4y}, & x>0,y>0\\ 0, & \text{其他}\end{cases}$则

$P\{X<Y\}=\iint\limits_{x<y}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{+\infty}\mathrm{d}y\int_0^{y}e^{-x-4y}\mathrm{d}x=\int_0^{+\infty}e^{-5y}\mathrm{d}y=\frac{1}{5}$,故选 A。

3-54 【2012 年数一,(8),4 分】将长度为 1m 的木棒随机地截成两段,则两段长度的相关系数为(　　)。

A. 1　　B. $\frac{1}{2}$　　C. $-\frac{1}{2}$　　D. -1

【分析】 设两段长度分别为X,Y,显然$X+Y=1$,即$Y=-X+1$,故两者是线性关系,且是负相关,所以相关系数为-1,故选 D。

3-55 【2013 年数三,(13),4 分】设随机变量X服从标准正态分布$X\sim N(0,1)$,则$E(Xe^{2X})=$________。

【解析】 标准正态分布的概率密度函数$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$$E(Xe^{2X})=\int_{-\infty}^{+\infty}xe^{2x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}xe^{-\frac{1}{2}(x-2)^2-2}\mathrm{d}x$$
$$=e^2\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-2)^2}\mathrm{d}x=2e^2$$

拓展实训Ⅱ

3-56 【2003 数四,十二题,13 分】对于任意二事件A和B,$0<P(A)<1$,$0<P(B)<1$,

$$\rho=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)P(B)P(\bar{A})P(\bar{B})}}$$

称做事件A和B的相关系数。

(1) 证明事件A和B独立的充分必要条件是其相关系数等于零;

(2) 利用随机变量相关系数的基本性质，证明 $|\rho|\leqslant 1$。

【分析】 (1) 利用事件 A 和 B 独立的定义 $P(AB)=P(A)P(B)$ 即可；

(2) 随机变量 X 和 Y 的相关系数为 $\rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\dfrac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$，

而需将 $\rho=\dfrac{P(AB)-P(A)P(B)}{\sqrt{P(A)P(B)P(\bar{A})P(\bar{B})}}$ 转化为用随机变量表示，显然，

若有 $E(XY)=P(AB)$，$E(X)=P(A)$，$E(Y)=P(B)$ 以及 $\sqrt{D(X)}=\sqrt{P(A)P(\bar{A})}$，$\sqrt{D(Y)}=\sqrt{P(B)P(\bar{B})}$ 即可，这只需定义

$$X=\begin{cases}1, & 若 A 出现\\ 0, & 若 A 不出现\end{cases},\quad Y=\begin{cases}1, & 若 B 出现\\ 0, & 若 B 不出现\end{cases}$$

【详解】 (1) 由 ρ 的定义，可见 $\rho=0$ 当且仅当

$$P(AB)-P(A)P(B)=0$$

而这恰好是二事件 A 和 B 独立的定义，即 $\rho=0$ 是 A 和 B 独立的充分必要条件。

(2) 考虑随机变量 X 和 Y。

$$X=\begin{cases}1, & 若 A 出现\\ 0, & 若 A 不出现\end{cases},\quad Y=\begin{cases}1, & 若 B 出现\\ 0, & 若 B 不出现\end{cases}$$

由条件知，X 和 Y 都服从 0-1 分布。

$$X\sim\begin{pmatrix}0 & 1\\ P(\bar{A}) & P(A)\end{pmatrix},\quad Y\sim\begin{pmatrix}0 & 1\\ P(\bar{B}) & P(B)\end{pmatrix}$$

易见

$$E(X)=P(A),\quad E(Y)=P(B)$$
$$D(X)=P(A)P(\bar{A}),\quad D(Y)=P(B)P(\bar{B})$$
$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=P(AB)-P(A)P(B)$$

因此，事件 A 和 B 的相关系数就是随机变量 X 和 Y 的相关系数。

于是由二随机变量相关系数的基本性质，可见 $|\rho|\leqslant 1$。

【评注】 如上 0-1 分布与随机事件之间的关系值得注意，较好地将两者联系起来，为借助相互的性质提供了便利。

3-57 【2004 数一，(22)，9 分】设 A,B 为随机事件，且 $P(A)=\dfrac{1}{4}$，$P(B|A)=\dfrac{1}{3}$，$P(A|B)=\dfrac{1}{2}$，令

$$X=\begin{cases}1, & A 发生\\ 0, & A 不发生\end{cases},\quad Y=\begin{cases}1, & B 发生\\ 0, & B 不发生\end{cases}$$

求：(1) 二维随机变量 (X,Y) 的概率分布；

(2) X 和 Y 的相关系数 ρ_{XY}。

【分析】 先确定 (X,Y) 的可能取值，再求在每一个可能取值点上的概率，而这可利用随机事件的运算性质得到，即得二维随机变量 (X,Y) 的概率分布；利用联合概率分布

可求出边缘概率分布，进而可计算出相关系数。

【详解】 (1) 因为 $P(AB)=P(A)P(B|A)=\frac{1}{12}$， 于是 $P(B)=\frac{P(AB)}{P(A|B)}=\frac{1}{6}$

则有 $P\{X=1,Y=1\}=P(AB)=\frac{1}{12}$

$$P\{X=1,Y=0\}=P(A\bar{B})=P(A)-P(AB)=\frac{1}{6}$$

$$P\{X=0,Y=1\}=P(\bar{A}B)=P(B)-P(AB)=\frac{1}{12}$$

$$P\{X=0,Y=0\}=P(\bar{A}\bar{B})=1-P(A\cup B)=1-[P(A)+P(B)-P(AB)]$$
$$=\frac{2}{3}$$

或 $P\{X=0,Y=0\}=1-\frac{1}{12}-\frac{1}{6}-\frac{1}{12}=\frac{2}{3}$

即 (X,Y) 的概率分布为

X \ Y	0	1
0	$\frac{2}{3}$	$\frac{1}{12}$
1	$\frac{1}{6}$	$\frac{1}{12}$

(2) 方法一：因为 $E(X)=P(A)=\frac{1}{4}$，$E(Y)=P(B)=\frac{1}{6}$，$E(XY)=\frac{1}{12}$

$E(X^2)=P(A)=\frac{1}{4}$， $E(Y^2)=P(B)=\frac{1}{6}$

$D(X)=E(X^2)-[E(X)]^2=\frac{3}{16}$， $D(Y)=E(Y^2)-[E(Y)]^2=\frac{5}{36}$

$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{24}$

所以 X 与 Y 的相关系数 $\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)D(Y)}}=\frac{1}{\sqrt{15}}=\frac{\sqrt{15}}{15}$

方法二：X,Y 的概率分布分别为

X	0	1
P	$\frac{3}{4}$	$\frac{1}{4}$

Y	0	1
P	$\frac{5}{6}$	$\frac{1}{6}$

则 $E(X)=\frac{1}{4}$，$E(Y)=\frac{1}{6}$，$D(X)=\frac{3}{16}$，$D(Y)=\frac{5}{36}$，$E(XY)=\frac{1}{12}$。

故 $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{24}$，从而

$$\rho_{XY}=\frac{\operatorname{Cov}(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}=\frac{\sqrt{15}}{15}$$

【评注】 本题尽管难度不大，但考查的知识点很多，综合性较强。通过随机事件定义随机变量或通过随机变量定义随机事件，可以比较好地将概率论的知识前后连贯起来。

3-58 【2004 数三，(22)；2004 数四，(22)，13 分】设 A,B 为两个随机事件，且 $P(A)=\frac{1}{4},P(B|A)=\frac{1}{3},P(A|B)=\frac{1}{2}$，令

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

求

(1) 二维随机变量 (X,Y) 的概率分布；

(2) X 与 Y 的相关系数 ρ_{XY}；

(3) $Z=X^2+Y^2$ 的概率分布。

【分析】 本题前面两个问题的解答见 2004 数一(22)，接下来仅就第(3)予以详解。

【详解】

(3) Z 的可能取值为 0,1,2。

$$P\{Z=0\}=P\{X=0,Y=0\}=\frac{2}{3}$$

$$P\{Z=1\}=P\{X=1,Y=0\}+P\{X=0,Y=1\}=\frac{1}{4}$$

$$P\{Z=2\}=P\{X=1,Y=1\}=\frac{1}{12}$$

即 Z 的概率分布为

Z	0	1	2
P	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{1}{12}$

【评注】 本题考查了二维离散随机变量联合概率分布，数字特征和二维离散随机变量函数的分布等计算问题，属于综合性题型。

3-59 【2006 数三，(22)；2006 数四，(23)，13 分】设随机变量 X 的概率密度为

$$f_X(x)=\begin{cases}\frac{1}{2}, & -1<x<0\\ \frac{1}{4}, & 0\leqslant x<2\\ 0, & \text{其他}\end{cases}$$

令 $Y=X^2$，$F(x,y)$ 为二维随机变量 (X,Y) 的分布函数。

(1) 求 Y 的概率密度 $f_Y(y)$；

(2) $\mathrm{Cov}(X,Y)$；

(3) $F\left(-\frac{1}{2},4\right)$。

【分析】 此题比 2006 数一(22)题多出(2)，其他(1)和(3)两问的答案请读者参见前面的解答。

【详解】 (1) 参看 2006 数一(22)解答。

(2) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(X,X^2)=E[X-E(X)][X^2-E(X^2)]=E(X^3)-E(X)E(X^2)$

而 $E(X)=\int_{-1}^{0}\frac{x}{2}\mathrm{d}x+\int_{0}^{2}\frac{x}{4}\mathrm{d}x=\frac{1}{4}$，$E(X^2)=\int_{-1}^{0}\frac{x^2}{2}\mathrm{d}x+\int_{0}^{2}\frac{x^2}{4}\mathrm{d}x=\frac{5}{6}$

$$E(X^3)=\int_{-1}^{0}\frac{x^3}{2}\mathrm{d}x+\int_{0}^{2}\frac{x^3}{4}\mathrm{d}x=\frac{7}{8}$$

所以 $\mathrm{Cov}(X,Y)=\frac{7}{8}-\frac{1}{4}\times\frac{5}{6}=\frac{2}{3}$。

(3) 参看 2006 数一(22)解答。

【评注】 本题属基本题型，只需注意计算的准确性，应该可以顺利求解。第一步求随机变量函数分布，一般都是通过定义用分布函数法讨论的。

3-60 【2006 数四，(22)，13 分】设二维随机变量 (X,Y) 的概率分布为

X \ Y	-1	0	1
-1	a	0	0.2
0	0.1	b	0.2
1	0	0.1	c

其中 a,b,c 为常数，且 X 的数学期望 $E(X)=-0.2$，$P\{Y\leqslant 0\mid X\leqslant 0\}=0.5$，记 $Z=X+Y$，

求

(1) a,b,c 的值；

(2) Z 的概率分布；

(3) $P\{X=Z\}$。

【分析】 利用二维离散型随机变量概率分布的性质和定义计算。

【详解】 (1) 由概率分布的性质知 $a+0.2+0.1+b+0.2+0.1+c=1$，即

$$a+b+c=0.4 \qquad ①$$

由 (X,Y) 可写出 X 的边缘概率分布为

X	-1	0	1
P	$a+0.2$	$b+0.3$	$c+0.1$

故 $E(X)=-(a+0.2)+(c+0.1)=-0.2$,即

$$a-c=0.1 \qquad ②$$

又因 $0.5=P\{Y\leqslant 0|X\leqslant 0\}=\dfrac{P\{X\leqslant 0,Y\leqslant 0\}}{P\{X\leqslant 0\}}=\dfrac{a+b+0.1}{a+b+0.5}$,即

$$a+b=0.3 \qquad ③$$

将①,②,③联立解方程组得

$$a=0.2,\quad b=0.1,\quad c=0.1$$

(2) Z 的可能取值为 $-2,-1,0,1,2$,则

$P\{Z=-2\}=P\{X+Y=-2\}=P\{X=-1,Y=-1\}=0.2$

$P\{Z=-1\}=P\{X=-1,Y=0\}+P\{X=0,Y=-1\}=0.1$

$P\{Z=0\}=P\{X=-1,Y=1\}+P\{X=0,Y=0\}+P\{X=1,Y=-1\}=0.3$

$P\{Z=1\}=P\{X=1,Y=0\}+P\{X=0,Y=1\}=0.3$

$P\{Z=2\}=P\{X=1,Y=1\}=0.1$

故 Z 的概率分布为

Z	-2	-1	0	1	2
P	0.2	0.1	0.3	0.3	0.1

(3) $P\{X=Z\}=P\{X=X+Y\}=P\{Y=0\}=0+0.1+0.1=0.2$

3-61 【2007 数三,(24),11 分】设随机变量 X 与 Y 独立同分布,且 X 的概率分布为

X	1	2
P	$\frac{2}{3}$	$\frac{1}{3}$

记 $U=\max(X,Y)$, $V=\min(X,Y)$

(1) 求 (U,V) 的概率分布;

(2) 求 U 与 V 的协方差 $\mathrm{Cov}(U,V)$。

【分析】 先写出 (U,V) 的可能取值,然后利用定义求概率。

【详解】 (1) (U,V) 的可能取值为 $(1,1),(1,2),(2,1),(2,2)$,则

$$P(U=1,V=1)=P(X=1,Y=1)=P(X=1)P(Y=1)=\frac{4}{9}$$

$$P(U=1,V=2)=0$$

$$\begin{aligned}P(U=2,V=1)&=P(X=2,Y=1)+P(X=1,Y=2)\\&=P(X=2)P(Y=1)+P(X=1)P(Y=2)=\frac{4}{9}\end{aligned}$$

$$P(U=2,V=2)=P(X=2,Y=2)=P(X=2)P(Y=2)=\frac{1}{9}$$

故 (U,V) 的概率分布为

V \ U	1	2
1	$\frac{4}{9}$	$\frac{4}{9}$
2	0	$\frac{1}{9}$

（2）由(U,V)的概率分布可得

$$E(U)=\frac{14}{9},\quad E(V)=\frac{10}{9},\quad E(UV)=\frac{16}{9}$$

所以 $\mathrm{Cov}(U,V)=E(U)E(V)-E(UV)=-\frac{4}{81}$

【评注】 本题为基础题型。

3-62 【2010年数三，(23)，11分】箱内有6个球，其中红，白，黑球的个数分别为1，2，3，现在从箱中随机地取出2个球，设X为取出的红球个数，Y为取出的白球个数，

（1）求随机变量(X,Y)的概率分布。

（2）求$\mathrm{Cov}(X,Y)$。

【分析】 （1）写出(X,Y)的可能值，然后计算相应的概率即得(X,Y)的概率分布；

（2）利用公式$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$计算。

【详解】 （1）X的可能值为0，1；Y的可能值为0，1，2，于是

$$P\{X=0,Y=0\}=\frac{C_3^2}{C_6^2}=\frac{3}{15},\quad P\{X=1,Y=0\}=\frac{C_3^1}{C_6^2}=\frac{3}{15},$$

$$P\{X=0,Y=1\}=\frac{C_3^1C_2^1}{C_6^2}=\frac{6}{15},\quad P\{X=1,Y=1\}=\frac{C_2^1}{C_6^2}=\frac{2}{15},$$

$$P\{X=0,Y=2\}=\frac{C_2^2}{C_6^2}=\frac{1}{15},\quad P\{X=1,Y=2\}=0$$

随机变量(X,Y)的概率分布为

X \ Y	0	1	2
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0

（2）由上可知X的概率分布和Y的概率分布分别为

X	0	1
P	$\frac{2}{3}$	$\frac{1}{3}$

Y	0	1	2
P	$\frac{6}{15}$	$\frac{8}{15}$	$\frac{1}{15}$

所以 $E(X)=\frac{1}{3}, E(Y)=\frac{2}{3}$。又 $E(XY)=\frac{2}{15}$，故

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{2}{15} - \frac{2}{3}\times\frac{1}{3} = -\frac{4}{45}$$

【评注】 本题求离散型二维随机变量的概率分布及两个随机变量的协方差，为基础题型。

3-63 【2011 年数一、数三，(22)，11 分】设随机变量 X 与 Y 的的概率分布分别为

X	0	1
P	1/3	2/3

Y	−1	0	1
P	1/3	1/3	1/3

且 $P\{X^2=Y^2\}=1$

求：(1) 求二维随机变量 (X,Y) 的概率分布；

(2) 求 $Z=XY$ 的概率分布；

(3) 求 X 与 Y 的相关系数 ρ_{XY}。

【考点分析】 本题考查二维离散型分布的分布律及相关数字特征的计算。其中，最主要的是第一问联合分布的计算。

【解析】 (1) 由于 $P\{X^2=Y^2\}=1$，因此 $P\{X^2\neq Y^2\}=0$。

故 $P\{X=0,Y=1\}=0$，因此

$P\{X=1,Y=1\}=P\{X=1,Y=1\}+P\{X=0,Y=1\}=P\{Y=1\}=1/3$

再由 $P\{X=1,Y=0\}=0$ 可知

$P\{X=0,Y=0\}=P\{X=1,Y=0\}+P\{X=0,Y=0\}=P\{Y=0\}=1/3$

同样，由 $P\{X=0,Y=-1\}=0$ 可知

$P\{X=0,Y=-1\}=P\{X=1,Y=-1\}+P\{X=0,Y=-1\}=P\{Y=-1\}=\frac{1}{3}$

这样，我们就可以写出 (X,Y) 的联合分布如下。

X \ Y	−1	0	1
0	0	1/3	0
1	1/3	0	1/3

(2) $Z=XY$ 可能的取值有 $-1,0,1$，

其中 $P\{Z=-1\}=P\{X=1,Y=-1\}=\frac{1}{3}$，　$P\{Z=1\}=P\{X=1,Y=1\}=\frac{1}{3}$，

则有 $P\{Z=0\}=\frac{1}{3}$。

因此，$Z=XY$ 的分布律为

Z	−1	0	1
P	1/3	1/3	1/3

(3) $E(X)=2/3, E(Y)=0, E(XY)=0, \mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0$

故 $\rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=0$

3-64 【2012 年数一、数三,(22),11 分】设二维离散型随机变量 X,Y 的概率分布为

X \ Y	0	1	2
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{3}$	0
2	$\frac{1}{12}$	0	$\frac{1}{12}$

求:(1) $P\{X=2Y\}$;

(2) $\mathrm{Cov}(X-Y,Y)$。

【解析】 (1) $P\{X=2Y\}=P\{X=0,Y=0\}+P\{X=2,Y=1\}=\frac{1}{4}+0=\frac{1}{4}$

(2) X 的概率分布为

X	0	1	2
P	1/2	1/3	1/6

故
$$E(X)=0\times\frac{1}{2}+1\times\frac{1}{3}+2\times\frac{1}{6}=\frac{2}{3}$$

XY 的概率分布为

XY	0	1	2	4
P	7/12	1/3	0	1/12

故
$$E(XY)=0\times\frac{7}{12}+1\times\frac{1}{3}+2\times0+4\times\frac{1}{12}=\frac{2}{3}$$

Y 的概率分布为

Y	0	1	2
P	1/3	1/3	1/3

$E(Y)=0\times\frac{1}{3}+1\times\frac{1}{3}+2\times\frac{1}{3}=1$

从而 $E(Y^2)=0^2\times\frac{1}{3}+1^2\times\frac{1}{3}+2^2\times\frac{1}{3}=\frac{5}{3}$

$D(Y)=E(Y^2)-[E(Y)]^2=\frac{5}{3}-1=\frac{2}{3}$

故 $\mathrm{Cov}(X-Y,Y)=\mathrm{Cov}(X,Y)-\mathrm{Cov}(Y,Y)=E(XY)-E(X)E(Y)-D(Y)=\frac{2}{3}-\frac{2}{3}\times 1-\frac{2}{3}=-\frac{2}{3}$。

3-65 【2012 年数三,(23),11 分】设随机变量 X 和 Y 相互独立,且均服从参数为 1 的指数分布,$V=\min(X,Y)$,$U=\max(X,Y)$。

求:(1) 随机变量 V 的概率密度;

(2) $E(U+V)$。

【解析】

(1) X 概率密度为 $f(x)=\begin{cases}\mathrm{e}^{-x}, & x>0\\ 0, & \text{其他}\end{cases}$,分布函数为 $F(x)=\begin{cases}1-\mathrm{e}^{-x}, & x>0\\ 0, & \text{其他}\end{cases}$,$X$ 和 Y 同分布。由 $V=\min(X,Y)$,$F_V(v)=P\{V\leqslant v\}=P\{\min(X,Y)\leqslant v\}=1-P\{X>v,Y>v\}$而 X,Y 独立,故上式等于 $1-P\{X>v\}P\{Y>v\}=1-[1-F(v)]^2=\begin{cases}1-\mathrm{e}^{-2v}, & v>0\\ 0, & \text{其他}\end{cases}$,故 $f_V(v)=F'_V(v)=\begin{cases}2\mathrm{e}^{-2v}, & v>0\\ 0, & \text{其他}\end{cases}$。

(2) 同理,U 的概率密度为 $f_U(u)=\begin{cases}2(1-\mathrm{e}^{-u})\mathrm{e}^{-u}, & u>0\\ 0, & \text{其他}\end{cases}$,

$$E(U)=\int_0^{+\infty}u2(1-\mathrm{e}^{-u})\mathrm{e}^{-u}\mathrm{d}u=\frac{3}{2},\quad E(V)=\int_0^{+\infty}v2\mathrm{e}^{-2v}\mathrm{d}v=\frac{1}{2}$$

所以
$$E(U+V)=E(U)+E(V)=\frac{3}{2}+\frac{1}{2}=2$$

3.5 应用实训

3-66 游戏中的公平。

国外有些游乐场里有一种赌博游戏,一个笼子里装了三粒骰子。把笼子摇一摇,停下来,三粒骰子各出现一个点数。参加游戏的人每次花一元钱买票,并且认定一个点数。例如,他认准“2”,如果有一粒骰子出现“2”,他就从游戏主持人那里赢回 1 元钱。运气好一点,两点骰子同时出现“2”,他赢回 2 元;三粒骰子都是“2”,赢回 3 元。同时,主持人还再退他 1 元钱。

这似乎是公平的游戏。如果 6 个参加者分别认定不同的 6 个点子,而骰子摇出“1”、“3”、“5”,那么主持人要向 6 人中的三人退还票钱,再各付 1 元。认定“2”,“4”,“8”的三人折赔票钱各 1 元,主持人收入 6 元,付出 6 元,而参加者三人赢,三人输,机会均等。

但是,有时两粒骰子出现相同的点数,主持人就只退给两人票钱。于是他收入 6 元而交出 5 元。当三粒骰子点数一样时,主持人收入 6 元而支出 4 元。这么一算,多数参加者总是要赔钱!

有的参加者不这么想，他觉得自己还是有利可图的"例如我认定'2'，掷一个骰子时，我赢的机会是$\frac{1}{6}$。可是现在是三粒骰子，我赢的机会是$\frac{1}{6}$的3倍，则$\frac{3}{6}=\frac{1}{2}$这是公平的！何况我还可能一次赢回2元、3元呢！"

那么，怎样才能准确地算出参加者平均每次的赢得钱数呢？

每玩一次游戏，参加者赢得的钱数x是不确定的量，如果三粒骰子都出现了他所要的点数，则他净赢3元，即$x=3$。这种情况在216($216=6\times6\times6$)个基本事件中只出现一次，故$x=3$的概率是$\frac{1}{216}$。

在216个基本事件中，有15种情形恰有两粒骰子出现所要的点数。这时他净赢2元，即$x=2$的概率是$\frac{15}{216}$。

类似地，$x=1$的概率是$\frac{75}{216}$。

最后，$x=-1$的概率是$\frac{125}{216}$。

即x的分布率为

x	-1	1	2	3
p	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$

$$E(x)=3\times\frac{1}{216}+2\times\frac{15}{216}+1\times\frac{75}{216}+(-1)\times\frac{125}{216}=\frac{-17}{216}$$

类似地，一个商店经理决定进一批羽绒服供应冬季市场。若今年冬天有寒流来袭，货将畅销，可获利2万元；若无寒流，气温正常，可获利1万元；若为暖冬，则将亏损5000元。根据历年气温记录与气象预报，估计有寒流的概率为$\frac{1}{6}$，正常的概率为$\frac{3}{4}$，暖冬的概率为$\frac{1}{12}$，于是获利的数学期望为

$$E(X)=2\times\frac{1}{6}+1\times\frac{3}{4}+(-0.5)\times\frac{1}{12}=\frac{25}{24}\approx1.04(\text{万元})$$

这表明，进一批羽绒服还是有赢利的希望的。

由此看来，我们可以通过求随机变量的数学期望迅速地对分布律的整体做出判断，例如，要比较不同班级的学习成绩，比较不同地区的粮食收成等。

3-67 企业经营决策。

(1) 灵活营销手段——提高利润。

在经济活动中，商业企业总是想方设法追逐更多的利润。为此，他们推出了各种名目繁多的活动，看似降低售价，让利于消费者，实质上还是为了提高利润。

某大型商场对某种原来售价2500元的家用电器进行"让利"促销活动，推出先使用后付款的方式。设该家用电器的使用寿命为X(单位：年)，规定：

$X\leqslant 1$	一台付款1500元	$1<X\leqslant 2$	一台付款2000元
$2<X\leqslant 3$	一台付款2500元	$X>3$	一台付款3000元

已知寿命 X 服从参数为1/10的指数分布，请估算该商场在促销活动中销售一台该家电利润是降低了还是提高了？

【解】 设 Y 为商家销售一台家电所获得的利润，则

$$Y=\begin{cases}1500, & X\leqslant 1\\ 2000, & 1<X\leqslant 2\\ 2500, & 2<X\leqslant 3\\ 3000, & X>3\end{cases}$$

$$P(X\leqslant 1)=\int_0^1 \frac{1}{10}e^{-\frac{1}{10}x}dx=F(1)=1-e^{-\frac{1}{10}}=0.095$$

$$P(1<X\leqslant 2)=\int_1^2 \frac{1}{10}e^{-\frac{1}{10}x}dx=F(2)-F(1)=e^{-\frac{1}{10}}-e^{-\frac{2}{10}}=0.086$$

$$P(2<X\leqslant 3)=\int_2^3 \frac{1}{10}e^{-\frac{1}{10}x}dx=F(3)-F(2)=e^{-\frac{2}{10}}-e^{-\frac{3}{10}}=0.078$$

$$P(X>3)=\int_3^{+\infty} \frac{1}{10}e^{-\frac{1}{10}x}dx=1-F(3)=e^{-\frac{3}{10}}=0.741$$

所以 Y 的分布律为

Y	1500	2000	2500	3000
$P(Y)$	0.095	0.086	0.078	0.741

故

$$\begin{aligned}E(Y)&=1500\times P(X\leqslant 1)+2000\times P(1<X\leqslant 2)+2500\\&\quad\times P(2<X\leqslant 3)+3000\times P(X>3)\\&=2732.5\end{aligned}$$

为此，需求出在促销活动中该电器售价 Y 的数学期望 $E(Y)$。先求出寿命 X 落在各时间区间内的概率，因为寿命 X 服从参数为1/10的指数分布，所以其概率密度就是 Y 的期望：2732.5元。由大数定律知，促销活动中该电器的平均售价约为2732元，每台电器利润提高了232元。

(2) 合理组织货源——获得最大经济利润。

如何获得最大利润是商界永远追求的目标，随机变量函数期望的应用为此问题的解决提供了新的思路。假定某商品在国际市场上每年对我国出口商品的需求量是随机变量 X(单位：t)，已知 X 服从[2000,4000]上的均匀分布，设每出售这种商品1t，可为国家挣得外汇3万元，但假如销售不出而囤积于仓库，则每吨需浪费保养费1万元，问应组织多少货源，才能使国家的平均收益最大？

【分析】 此问题的解决先是建立利润与需求量的函数，然后求利润的期望，从而得到利润关于货源的函数，最后利用求极值的方法得到答案。

【解】 Y：每年该商品的出口量，R：收益，X：需求量。

$$f(x)=\begin{cases}\dfrac{1}{2000}, & 2000\leqslant x\leqslant 4000\\ 0, & \text{其他}\end{cases}$$

$$R(x)=\begin{cases}3y, & x\geqslant y\\ 3x-(y-x), & x\leqslant y\end{cases}\quad x,y\in[2000,4000]$$

$$E(R)=\int_{-\infty}^{+\infty}R(x)f(x)\mathrm{d}x=\int_{2000}^{y}(4x-y)\frac{1}{2000}\mathrm{d}x+\int_{y}^{4000}3y\frac{1}{2000}\mathrm{d}x$$

$$=\frac{1}{1000}(-y^2+7000y-4\times10^6)=\frac{1}{1000}[825\,000-(y-3500)^2]$$

上述计算表明 $E(R)$是 y 的二次函数，通常用求极值的方法可以求得，$y=3500\text{t}$ 时，能够使得期望的利润达到最大。

(3) 恰当选择投资领域——有效规避风险。

在进行经济管理决策之前，往往存在不确定的随机因素，从而所作的决策有一定的风险。只有正确、科学的决策才能达到以最小的成本获得最大的安全保障的总目标，才能尽可能节约成本。利用概率统计知识可以获得合理的决策，从而实现这个目标。下面以数学期望、方差等数字特征为例说明它在经济管理决策中的应用。

某人有一笔资金，可投入三个项目：房产 x、地产 y 和商业 z，其收益和市场状态有关，若把未来市场划分为好、中、差三个等级，其发生的概率分别为 $p_1=0.2$，$p_2=0.7$，$p_3=0.1$，根据市场调研的情况可知不同等级状态下各种投资的年收益(万元)，见表 3.1。

表 3.1

	好 $p_1=0.2$	中 $p_2=0.7$	差 $p_3=0.1$
房产	11	3	−3
地产	6	4	−1
商业	10	2	−2

请问：该投资者如何投资好？

【解】 我们先考查数学期望，可知

$$E(x)=11\times0.2+3\times0.7+(-3)\times0.1=4.0$$
$$E(y)=6\times0.2+4\times0.7+(-1)\times0.1=3.9$$
$$E(z)=10\times0.2+2\times0.7+(-2)\times0.1=3.2$$

根据数学期望可知，投资房产的平均收益最大，可能选择房产，但投资也要考虑风险，我们再来考虑它们的方差。

$$D(x)=(11-4)^2\times0.2+(3-4)^2\times0.7+(-3-4)^2\times0.1=15.4$$
$$D(y)=(6-3.9)^2\times0.2+(4-3.9)^2\times0.7+(-1-3.9)^2\times0.1=3.29$$
$$D(z)=(10-3.2)^2\times0.2+(2-3.2)^2\times0.7+(-2-3.2)^2\times0.1=12.96$$

因为方差越大，则收益的波动越大，从而风险也越大，所以从方差看，投资房产的风险比投资地产的风险大得多，若收益与风险综合权衡，该投资者还是应该选择投资地产为

好，虽然平均收益少 0.1 万元，但风险要小一半以上。

(4) 选择最佳设备引入方式——节约成本投入。

某高级镜片制造厂试制成功新镜头，准备出口试销，厂方自制的检测设备与国外的检测设备仍有一定的差距，为此，厂方面临一个决策问题：①直接进口，②租用设备，③与外商合资。不同的经营方式所需的固定成本和每件的可变成本如下表所示。

	自制	进口	租赁	合资
固定成本/万元	120	40	64	200
每件可变成本/元	60	100	80	40

已知产品出口价为 200 元/件，如果畅销可销 3.5 万件，中等可销 2.5 万件，滞销只售 0.8 万件，按以往经验，畅销的可能性为 0.2，中等的为 0.7，滞销的为 0.1，请为该厂做出最优决策。

【解】 设 B＝销量，A_1＝自制，A_2＝进口，A_3＝租赁，A_4＝合资。

$E(B)=2.53$ 万件

$$E(A_1)=2.53\times 200-(120+2.53\times 60)=234.2$$
$$E(A_2)=2.53\times 200-(40+2.53\times 100)=213$$
$$E(A_3)=2.53\times 200-(64+2.53\times 80)=239.6$$
$$E(A_4)=2.53\times 200-(200+2.53\times 40)=204.8$$

所以 A_3 为最优方案，即租用设备。

3-68 企业风险管理。

(1) 风险评价。

企业面临多种风险，在一定条件下企业侧重于其中很关键的某一风险或某一部分风险管理。如果已知某一风险事故发生所导致损失的概率分布就能算出风险的期望值 μ 与方差 σ^2 及标准差 σ，期望值是平均受损额，方差与标准差都是显示风险损失的变动幅度的。标准差越大风险就越难把握，但是其平均受损额很大时一般的损失变动可以认为相对风险不大。真实反映风险大小的量是差异系数 σ/μ，如表 3.2 所示。

表　3.2

某超市雨天损失的概率分布表					
损失金额/万元	1.5	2.8	3.6	3.9	4.1
概率	0.07	0.18	0.35	0.24	0.16
某超市高温天气损失的概率分布表					
损失金额/万元	0.6	0.8	1.1	1.5	2.3
概率	0.15	0.2	0.35	0.25	0.05

由雨天造成的损失分布表得

① 期望值。

$$\begin{aligned}\mu &= \sum_{i=1}^{5} x_i p_i \\ &= 1.5\times 0.07 + 2.8\times 0.18 + 3.6\times 0.35 + 3.9\times 0.24 + 4.1\times 0.16 \\ &= 3.461\end{aligned}$$

② 方差。

$$\begin{aligned}\sigma^2 = \sum_{i=1}^{5} (x_i-\mu)^2 p &= (1.5-3.461)^2\times 0.07 + (2.8-3.461)^2\times 0.18 \\ &\quad + (3.6-3.461)^2\times 0.35 + (3.9-3.461)^2\times 0.24 \\ &\quad + (4.1-3.461)^2\times 0.16 = 0.4662\end{aligned}$$

标准差 $\sigma=0.6827$(万元),差异系数 $V=\sigma/\mu=0.6827/3.461=0.1973$。

同理可得因高温天气损失概率分布数据为

期望值 $\mu=1.125$(万元),标准差 $\sigma^2=0.4085$(万元),差异系数 $V=0.3631$。

高温天气损失的差异系数大于雨天损失的差异系数,差异系数越大其风险也就越大,说明该超市因高温天气引起的损失风险大于雨天的损失风险。

(2) 风险决策。

决策有一定的风险,所以称之为风险型决策。在进行风险决策时,面临不同的客观状态,根据不同的状态会有几种行动方案可供选择。人们不能选择客观状态,但可以选择不同的方案。对于如何选择一个最佳的行动方案,就是概率论与数理统计要解决的问题。一般有以下三种常用的风险决策方法。

方法一:收益期望最大决策法。

收益期望最大决策法就是在各种行动方案中选择收益期望最大的方案作为最优方案,如表 3.3 所示。

表 3.3

某商场签字笔销售量的概率分布表

需求量	100	200	300	400	500	600
概率	0.05	0.1	0.26	0.35	0.15	0.1

当签字笔的进货量是 300 支时,如果市场需求为 100 支,则可以销售 100 支,每支笔卖 6 元,则可获利 600 元,还剩 200 支作为降价处理,每支笔卖 3 元,则共亏损 600 元。此时获利为零;市场需求为 200 支时,获利 1200 支减去亏损的 300 支得 900 支;市场需求为 300 支或大于 300 支时,获利 1800 元。

当进货量为 300 支时收益的期望值为

$$\begin{aligned}&0\times 0.05 + 900\times 0.10 + 1800\times 0.25 + 1800\times 0.35 \\ &+ 1800\times 0.15 + 1800\times 0.10 = 1620(\text{元})\end{aligned}$$

当进货量为 400 支时收益额期望值为

$$\begin{aligned}&-300\times 0.05 + 600\times 0.10 + 1800\times 0.25 + 1500\times 0.35 + 2400 \\ &\times 0.35 + 2400\times 0.15 + 2400\times 0.10 = 1860(\text{元})\end{aligned}$$

同理得：进货量分别在 100 支、200 支、500 支和 600 支时的收益分别为 600 元、1155 元、1785 元和 1575 元。

所以当签字笔的进货量在 400 支时收益的期望值为 1860 元。从长期来看，每年进货量为 400 支，则每年获得的平均利润为 1860 元。它是所有方案中期望获利最大的，可作为最佳方案。

方法二：损失期望值最小决策法。

损失期望值最小决策法就是在各种行动方案中选择一个损失的期望最小的方案作为最优方案。要使销售年利润最大，也就是使各种损失最小。表 3.3 中的数据表明，销售签字笔的损失费有两部分。一部分是签字笔的进货量大于需求时，因滞销而做降价处理；另一部分是当签字笔的进货量小于需求时，因缺货而失去销售机会的损失。总损失费为滞销损失费或缺货损失费。当决策方案为进货 300 支时，市场需求 100 支，则有 200 支要做降价处理，每支亏 3 元，出现滞销损失费 600 元；若市场需要 300 支时，此时总损失费为零。若市场需求量分别为 400 支、500 支、600 支时，出现缺货损失，损失费分别为 600 元、1200 元、1800 元，此时没有滞销损失。

进货 300 支时，损失费的期望值为

$$600\times 0.05+300\times 0.10+0\times 0.25+600\times 0.35$$
$$+1200\times 0.15+1800\times 0.10=630(\text{元})$$

同理可算得损失费期望值分别为 1650 元、1950 元、630 元、390 元、465 元、675 元。比较各损失费的期望值可知，当进货量为 400 支时，损失费的期望值最小。每年进货 400 支，长期来看每年平均总损失费最小为 390 元，故为最佳方案。

方法三：等概率决策法。

等概率决策法就是根据各种行动方案和各种自然状态的概率相等，计算可行方案作为最佳方案的收益期望值，选取收益期望最大的方案作为最佳方案的方法。当客观事物的结果不确定或信息不足时，不能确定哪种自然状态出现的可能性的大小时，就应假定各种自然状态发生的可能性是一样的，即概率大小是相等的。按此概率分布计算收益期望值并比较大小，选择期望值最大的行动方案为最优方案。因为是等概率事件，所以在表 3.3 中需求量在 100 支、200 支、300 支、400 支、500 支、600 支时的概率都是 1/6。

当签字笔的进货量是 300 支时，如果市场需求为 100 支，则可以销售 100 支，每支笔卖 6 元，则获利 600 元，还剩 200 支作为降价处理，每支价格 3 元，共亏损 600 元，此时获利是零；市场需求是 200 支时，获利 1200 元亏损 300 元得 900 元；市场需求为 300 支和 300 支以上时，获利 1800 元。

当进货量为 300 支时收益的期望值为

$$(0+900+1800+1800+1800+1800)\times 1/6=1350(\text{元})$$

同理得：进货量分别为 100 支、200 支、400 支、500 支、600 支时的收益为 600 元、1050 元、1500 元、1350 元。所以当签字笔的进货量在 400 支和 500 支时，收益期望值为 1500 元，是所有方案中期望获利的最大值，仍可选择进货 400 支为最佳方案。

3-69　化验中的节约。

人类的活动会使大量的工业、农业和生活废弃物排入水中，使水受到污染。目前，全

世界每年约有 4200 多亿立方米的污水排入江河湖海，污染了 5.5 万亿立方米的淡水，这相当于全球径流总量的 14%以上。

中国有 82%的人饮用浅井和江河水，其中水质污染严重、细菌超过卫生标准的占 75%，受到有机物污染的饮用水人口约 1.6 亿。长期以来，人们一直认为自来水是安全卫生的。但是，因为水污染，如今的自来水已不能算是卫生的。一项调查显示，在全世界自来水中，测出的化学污染物有 2221 种之多，其中有些已确认为致癌物或促癌物。要确定水源的安全性，就要对来自不同水源的水进行检验，确定是否被污染。如果对所有的水源逐一检验，这无疑是一项巨大的工作。尽管自来水被检出污染的比例很低，但这个检验是决定一个地区居民饮水是否安全的关键。如何保证“有污染的”会被检出，同时又减少检验次数呢？

假设被检水源中平均 20 个中有一个被污染，也就是说，将 20 个水源分为一组，对每一组进行 20 次检验，则平均每一组有一例呈阳性。显然，如果把几个水源混合起来进行检查，仅当至少有一个水源被污染时，污染物被检出。代替 20 次单个检验，我们把 20 个分为两组，对 10 个一组的两个混合水源样本分别进行检验。平均来说，此时一个混合样本检出污染，另一个未检出。然后仅对检出污染的混合样本进行单个检验，以确认哪一个水源是含有污染的。这样，对每 20 个一组平均仅需 2＋10＝12 次检验，即减少了 20 次中的 8 次，或减少 40%。可以看到，如果把 20 个样本按 5 个一组进行混合，则平均实验总数仅有 4＋5＝9 次，这是对 20 个申请者一组进行检验所需次数的最佳值，减少了 11 次，即 55%。

怎么达到上述最佳值呢，下面我们把问题一般化，假设水被污染的概率为 p，且各地水源被污染是相互独立的，现将 k 个水源混合为一组，设被检水源总数 n 为 k 的倍数，第 i 组需化验次数为 X_i，则 X_i 的分布律为

X_i	1	$k+1$
p_k	$(1-p)k$	$1-(1-p)k$

$$E(X_i)=(1-p)^k+(k+1)[1-(1-p)^k]=(k+1)-k(1-p)^k$$

$$E(X)=\sum_{i=1}^{n/k}E(X_i)=\frac{n}{k}[(k+1)-k(1-p)^k]=n\left[1-\left((1-p)^k-\frac{1}{k}\right)\right]$$

求最佳值过程依赖于要调查水源的污染率。即我们只要根据具体问题，找到使得上式达到最小的 k，就可得到最优的化验设计。

这个思想非常漂亮，混合样本监测的方法现已广泛实践于环境保护研究和其他领域，用于削减实验检测费用。

3-70 合理制定“犯法成本”，使违法商家无利可图。

近几年 3·15 陆续曝光了各类产品质量问题，包括食品安全、卫生医疗、通信、电子设备等，由于与人民群众息息相关，无疑是生活中一大热点。从食品中的三聚氰氨奶粉、苏丹红鸡蛋、地沟油火锅、瘦肉精火腿等，到用品中的垃圾变身餐巾纸、“百毒缠身”卫生筷等，再到赖于“安居”的“楼歪歪”、“窗飘飘”等。我们不禁要问，对于违法商家罚款到底是

多少时，他们才无利可图？又如何从罚款额度和查处力度来加强监管呢？

假设不法商家出售第 i 批不合格食品非法所得为 n 万元，$i=1,2,\cdots,m$，m 的取值由查处时出售的总批次确定，按质检部门陈述，一般在节假日会加大质量检查力度，则根据今年国家法定节假日 29 天和 52 个周末中只可能是 2 天中的任意 1 天来抽查计算，重点抽查概率仅为 $\frac{81}{365}=0.2219$，罚款以 y 万元计算，违法商家可以获得的平均利润必定远大于 0，即 $E(X_i)=0.7781n-y0.2219\geqslant 0$，则 $y\geqslant 3.5065n$ 万元，即每一批不合格产品至少罚款 3.5065n 万元，不法商家才无利可图，如果一直没有被查处，则利润为 nm 万元，要远远大于 50 万元是非常容易的，这就是问题所在了，“犯法成本”太低，而利润大得诱人。可以说即使是最高罚款 50 万元，与非法所得相比也仅仅是九牛一毛。这就是为什么在我国接连不断地出现食品安全事件的原因，与相关法律的惩罚力度有必然联系，丝毫不能威慑到不法企业。由上面的计算可知，在罚款额度确定的情况下，如果日常抽查概率提高，则单批违法所得的 $E(X_i)$ 将大幅减少，不法商家的数量必将随之快速减少，呈正比例关系。

3-71　均值与方差。

张家有钱一千万，
九个邻居穷光蛋。
平均起来算一算，
个个都是张百万。

这是互联网上的一首打油诗。诗中讲到九个穷光蛋邻居因为与“张千万”一起算平均财产，不幸“被富人”了，都成了“百万”级的富翁。这种说法显然有些夸张，但是它揭示了一个道理，即仅靠平均值来了解一个群体的收入情况是不妥的，它有时会掩盖严重的分配不均的事实。

“平均值”是统计学里的一个基本概念。你可别小看这个不起眼的东西，虽然它平凡得犹如你的口头禅一样，但弄不清楚它可能还要出人命呢。

传说很久以前某诸侯国的一次战争中，一位将军带领部队要渡过一条大河去执行作战任务。河上没有桥梁也没有任何船只，而将军的多数士兵都是旱鸭子，不会游泳。将军便派他的军师去调查一下河水的深浅。这位军师非常认真地提供了具体的信息。随后，军师向将军报告：“这里的老乡说，这条河平均水深是 1.4m。”将军问：那我们的士兵的平均身高是多少呢？军师回答：“1.65m。”将军面露微笑：“太好了，水深 1.4m，身高1.65m，士兵们的头正好可以露在水面上。”于是，将军下令：“全体官兵继续前进，涉水过河。”结果怎么样呢？很快就传来了士兵被淹死的消息。你说这是为什么呢？

显然，这条河各处的水深并不是一样的，老乡们说河的平均水深是 1.4m，是将各处深浅不同的河水深度平均后得到的。可以肯定地说，有的地方比较浅，不足 1.4m，有的地方比较深，则超过 1.4m，以致超过一些士兵的身高，若河水深度达到 1.8m，那些身高不足 1.8m 的士兵们很可能就会被淹死。同样，士兵的平均身高是 1.65m，也是将所有士兵的身高平均后得到的，士兵中有的身高比平均值大，有的身高比平均值小。若士兵身高是

1.5m，很可能还没有走到河水最深处就已经被淹死了。呵呵，真能出人命吧！没有一点统计学的常识，不懂平均值的概念多危险啊！

在现实生活中，我们一般关心平均数，由上面的例子，我们可以看到，有时平均数是会骗人的。举一个简单的例子，由5人组成的家庭篮球队，平均年龄23岁，该是一帮生龙活虎的小伙子吧？上场一看，原来是一位七旬老人领着4个十一二岁的娃娃！

可见，平均数不一定能代表典型的情况。知道了平均年龄是23岁，还应当看看几个人的年龄是不是接近平均年龄。如果5个人年龄分别是70、12、11、11、11，考查它们与平均年龄23的差的绝对值。$|70-23|$，$|12-23|$，$|11-23|$，$|11-23|$，$|11-23|$，把这几个数加起来再求平均值，就能反映出具体各人年龄与平均年龄的差异有多大。但是在数学上，更方便的是把这5个差的平方加起来求平均值，这个平均值叫做方差，即

$$\frac{(70-23)^2+(12-23)^2+(11-23)^2+(11-23)^2+(11-23)^2}{5}=552.4$$

3-72 体育竞技状态评价。

众所周知，在竞技体育比赛中，通过统计数字可以很好地反映一名运动员或是一支运动队在各方面的情况。下面就以NBA为例来讨论一下期望与标准差在竞技体育中的应用。

作为全球顶级的职业篮球联赛，NBA除了为广大球迷推出一道道明星荟萃的PK盛宴外，也巨细无遗地留下了海量的技术统计资料，诸如得分、篮板、助攻、胜率等技术指标令人目不暇接，眼花缭乱。在这个数字的茫茫大海中，难道真的是杂乱无章、毫无规律可循的吗？其实不然，就篮球这项运动的本质而言，从统计科学的角度来看无非是一种概率的集体博弈，从比赛双方的每一次进攻或防守，到球队的每一次选秀或交易，甚至是球员的每一次伤病，都可以看作一次随机事件，因此涉及的种种技术指标也就成为了随机变量。作为统计学中最重要的概率分布规律，正态分布就像一只无形的魔棒，操纵着NBA的方方面面。接下来看看正态分布到底是如何对NBA施展它的魔法的。

NBA球员某项技术指标的稳定性是由该技术指标分布的标准差决定的，这个值越小，那么他的这项技术指标就越稳定。

以姚明在三个赛季常规赛每场比赛中的得分分布为例，如下所示。

赛季	平均值	标准差
2003—2004	17.54	6.901
2004—2005	18.34	6.801
2005—2006	19.91	6.309

可见姚明2005—2006赛季在得分的稳定性方面有明显的进步，因为他在比赛中得到20分左右的概率增大了，而拿10分以下或拿30分以上的概率则相应地减少了，不再大起大落，这就是稳定性的体现。

再来看看一些NBA著名球星每个赛季常规赛场均得分的分布情况，如下所示。

球星	平均值	标准差
迈克尔·乔丹	30.7	3.72
哈基姆·奥拉朱旺	21.0	5.98
蒂姆·邓肯	22.2	1.62
科比·布莱恩特	23.4	7.62
特雷西·迈克格雷迪	21.8	8.89
沙克·奥尼尔	26.1	3.25
史蒂夫·弗朗西斯	19.3	2.11
文斯·卡特	23.1	4.09
凯文·加内特	20.4	3.90
阿伦·艾弗森	28.1	3.57

由此可以看出，邓肯的稳定性令人惊叹，无愧于“石佛”的称号；虽然科比和迈克格雷迪均为得分高手，但他们与乔天王相比还有较大的差距，或许伤病是造成这种情况的一大原因吧；在中锋这个位置上，奥尼尔的稳定性相当突出，说明他的竞技状态保持得比较好。

3-73 重回帽子问题。

参加 party 的 n 个人将他们的帽子放在一起，结束后每人任取一顶帽子戴上，试求戴对自己的帽子的人数 X 的均值和方差。

【解】 记 $X_i=\begin{cases}1, & \text{第 } i \text{ 个人戴对}\\ 0, & \text{第 } i \text{ 个人未戴对}\end{cases}\quad i=1,2,\cdots,n$

则 $X_1,X_2,\cdots,X_n$ 是同分布的，但不独立，其共同分布为

$$P(X_i=1)=\frac{1}{n},\quad P(X_i=0)=1-\frac{1}{n},\quad i=1,2,\cdots,n$$

由此得

$$E(X_i)=\frac{1}{n},\quad D(X_i)=\frac{1}{n}\left(1-\frac{1}{n}\right),\quad i=1,2,\cdots,n$$

又因为 $X=X_1+X_2+\cdots+X_n$，所以

$$E(X)=E(X_1)+E(X_2)+\cdots+E(X_n)=n\cdot\frac{1}{n}=1$$

但因为 X_i 间不独立，所以

$$D(X)=\sum_{i=1}^{n}D(X_i)+2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{Cov}(X_i,X_j)$$

为计算 $\mathrm{Cov}(X_i,X_j)$，先给出 X_iX_j 的分布列，注意到 X_iX_j 的可能取值为 0.1，且

$$P(X_iX_j=1)=P(X_i=1,X_j=1)=\frac{1}{n}\cdot\frac{1}{n-1}$$

所以

$$E(X_iX_j)=0\times P(X_iX_j=0)+1\times P(X_iX_j=1)=\frac{1}{n(n-1)}$$

因此

$$\mathrm{Cov}(X_i,X_j)=E(X_iX_j)-E(X_i)E(X_j)=\frac{1}{n(n-1)}-\left(\frac{1}{n}\right)^2=\frac{1}{n^2(n-1)}$$

由此得

$$D(X)=\frac{n-1}{n}+2C_n^2\frac{1}{n^2(n-1)}=1$$

3-74 中心极限定理在商业管理中的应用。

(1) 保险问题。

① 保险盈亏。

大数定律和中心极限定理是近代保险业赖以建立的基础。一个保险公司的盈亏,我们通过学习中心极限定理的知识都可以做到估算和预测。下面以一保险业的实例来阐述大数定律和中心极限定理在保险业中的重要作用。

某市保险公司开办一年期大病医疗保险业务,被保险人每年交付保险费100元,若一年内发生重大人身事故,则其本人或家属可获1万元赔金。已知该市人员一年内发生重大人身事故的概率为0.005,现有5000人参加此项保险,问保险公司一年内从此项业务所得总收益为20万~40万的概率及保险公司亏本的概率。

这一年中发生重大事故的人数为X,发生事故的概率为$p=0.005$,$X\sim B(5000,0.005)$,则$np=25$,$np(1-p)=24.875$,总收益为$0.01\times5000-X=50-X$,由中心极限定理得

$$P(20\leqslant 50-X\leqslant 40)=P(10\leqslant X\leqslant 30)=P\left(\frac{10-25}{\sqrt{24.875}}\leqslant\frac{X-25}{\sqrt{24.875}}\leqslant\frac{30-25}{\sqrt{24.875}}\right)$$
$$=\Phi(1.003)-\Phi(-3.008)=0.8403$$

而保险公司亏本的概率为

$$P(50<X)=P\left(\frac{X-25}{\sqrt{24.875}}>\frac{50-25}{\sqrt{24.875}}\right)$$
$$=1-\Phi(5.013)=1-0.999\,999\,830\,2$$
$$=0.000\,000\,169\,8$$

可见,保险行业只要发生意外事故的概率足够低(本例仅为0.005),这样的业务几乎是稳赚不赔的,这也是为什么保险行业乐于开展此类业务的原因。当然,作为投保人,“一保在手,有惊无险”,这对于我们抵御意外事故是非常有保障的。

② 保险中的大数。

保费的确定是要合理的,不可随便波动,而且保险公司要保证财政稳定,需要大量的客户,那么需要的“大数”具体为多少呢?大数定律告诉我们,当被保单位足够多时,实际损失和预期损失之间的误差会变得很小很小。

具体发生的概率和大数的大小需要利用大数定律和中心极限定理来计算出来。

设$X_1,X_2,X_3,\cdots,X_n$是独立同分布的随机变量,$E(X_i)<+\infty$,$D(X_i)<+\infty(i=1,2,\cdots,n)$有

$$\lim_{n\to\infty}\left(\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}E(X_i)}{\sqrt{\sum_{i=1}^{n}D(X_i)}}<x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{t^2}{2}}\mathrm{d}t$$

这里$X_1,X_2,X_3,\cdots,X_n$看作n个被保险人发生损失的次数,每个X_i取值为0或者1,

则 $X = \sum_{i=1}^{n} X_i$ 为 n 个被保险人发生损失的总人数，服从二项分布，当 n 很大时就可以计算出一个范围内人数发生损失的概率。例如，当 n 很大时，假设每个人发生损失的概率为 p，要求损失人数在 (l,m)（$l < m$，且都是正整数）(l,m) 的概率，那么就有

$$P(1 < x < m) \approx \Phi\left[\frac{m-np}{\sqrt{np(1-p)}}\right] - \Phi\left[\frac{l-np}{\sqrt{np(1-p)}}\right]$$

例如 10 000 个人中每个人死亡的概率都是 0.003，求死亡人数为 150～300 的概率，那么就有

$$P(150 < x < 300) \approx \Phi\left[\frac{300-30}{\sqrt{30(0.997)}}\right] - \Phi\left[\frac{150-30}{\sqrt{30(0.997)}}\right]$$

这样看来，保险需要大数估计，那么就需要足够多的人来参与买保险，实际损失和预期损失之间的误差才会很小。要如何确定这个“大数”，才能使得保险业得以存在和发展呢？

当随机变量服从 $X \sim N(\mu, \sigma^2)$ 时，$\bar{X} \sim N\left(p, \frac{p(1-p)}{n}\right)$，从而 $\bar{X}$ 的一个置信水平为 $1-\alpha$ 的置信区间为 $\left[p - \frac{\sqrt{p(1-p)}}{n} Z_{\frac{\alpha}{2}}, p + \frac{\sqrt{p(1-p)}}{n} Z_{\frac{\alpha}{2}}\right]$，实际损失变动与保单总数的比率为 $\varepsilon = \frac{\sqrt{p(1-p)}}{n^2} Z_{\frac{\alpha}{2}}$，若需要 ε 小于某个具体常数 k，则可由 $\varepsilon \leqslant k$ 解出 $n \geqslant \sqrt{\frac{\sqrt{p(1-p)}}{k} Z_{\frac{\alpha}{2}}}$，即可确定最低保单个数。

例如：20 岁时人的死亡率为 $p = 0.002$，某一寿险公司希望有 95% 的把握估计实际死亡人数能在预期死亡率 0.002 上下摆动的幅度不超过 $10\% \times 0.002$，那么这个公司需要承保人数必须达到

$$N = \frac{2^2 \times 0.002 \times 0.998}{0.0002^2} = 199\,600$$

（2）水房拥挤问题。

假设某高校有学生 5000 人，只有一个开水房，由于每天傍晚打开水的人较多，经常出现同学排长队的现象，为此校学生会特向学校后勤集团公司提议增设水龙头。假设后勤集团公司经过调查，发现每个学生在傍晚一般有 1% 的时间要占用一个水龙头，现有水龙头数量为 45 个，现在总务处遇到的问题是：

① 未新装水龙头前，拥挤的概率是多少？

② 需至少要装多少个水龙头，才能以 95% 以上的概率保证不拥挤？

【解】 ① 设同一时刻，5000 个学生中占用水龙头的人数为 X，则

$$X \sim B(5000, 0.01)$$

拥挤的概率是

$$P(X > 45) = 1 - P(0 \leqslant X \leqslant 45) = 1 - \sum_{k=0}^{45} C_{5000}^{k} \times 0.01^k \times 0.99^{5000-k}$$

直接计算相当麻烦，我们利用棣莫佛-拉普拉斯定理。已知 $n = 5000$，$p = 0.01$，$q = 0.99$，$np = 50$，$\sqrt{npq} = 7.04$。

故

$$P(0 \leqslant X \leqslant 45) \approx \Phi\left(\frac{45-50}{7.04}\right)-\Phi\left(\frac{0-50}{7.04}\right)=\Phi(-0.71)-\Phi(-7.1)=0.2389$$

从而 $p(\xi>45)=1-0.2389=0.761$，怪不得同学们有不少的抱怨，拥挤的概率竟达到了 76.11%。

② 欲求 m，使得 $P(0\leqslant X\leqslant 45)\geqslant 0.95$。

即 $$\Phi\left(\frac{m-50}{7.04}\right)-\Phi\left(\frac{0-50}{7.04}\right)\geqslant 0.95$$

由于 $$\Phi\left(\frac{0-50}{7.04}\right)=\Phi(-7.09)\approx 0$$

即 $$\Phi\left(\frac{m-50}{7.04}\right)\geqslant 0.95$$

查标准正态分布表，得 $$\frac{m-50}{7.04}\geqslant 1.645$$

即 $$m\geqslant 61.6$$

故需要装 62 个水龙头。

问题的变形：③需至少安装多少个水龙头，才能以 99%以上的概率保证不拥挤？欲求 m，使得 $P(0\leqslant X\leqslant 45)\geqslant 0.99$，

即 $$\Phi\left(\frac{m-50}{7.04}\right)-\Phi\left(\frac{0-50}{7.04}\right)\geqslant 0.99$$

由于 $$\Phi\left(\frac{0-50}{7.04}\right)=\Phi(-7.09)\approx 0.76$$

即 $$\Phi\left(\frac{m-50}{7.04}\right)\geqslant 0.99$$

查标准正态分布表，得 $$\frac{m-50}{7.04}\geqslant 2.325$$

即 $$m\geqslant 66.4$$

故需要装 67 个水龙头。

④ 若条件中已有水龙头数量改为 55 个，其余的条件不变，①，②两问题结果如何？

$$P(X>55)=1-\Phi\left(\frac{55-50}{7.04}\right)=1-\Phi(0.71)=0.2389$$

⑤ 若条件中的每个学生占用时间由 1%提高到 1.5%，其余条件不变，则①，②两问题结果如何？

【解】 ① 设同一时刻，5000 个学生中占用水龙头的人数为 X，则 $X\sim B(5000,0.015)$。

已知 $n=5000, p=0.015, q=0.985\ np=75, \sqrt{npq}=8.60$。

拥挤的概率是 $P(X>45)=1-\Phi\left(\frac{45-75}{8.60}\right)=1-\Phi(-3.49)\approx 1$。

拥挤的概率竟达到 100%。

② 欲求 m，使得 $P(0\leqslant X\leqslant 45)\geqslant 0.95$，

即 $$\Phi\left(\frac{m-75}{8.60}\right)-\Phi\left(\frac{0-75}{8.60}\right)\geqslant 0.95$$

由于 $$\Phi\left(\frac{0-75}{8.60}\right)\approx 0$$

即
$$\Phi\left(\frac{m-75}{8.60}\right)\geqslant 0.95$$
查标准正态分布表，得
$$\frac{m-75}{8.60}\geqslant 1.645$$
即
$$m\geqslant 89.14$$
故需要装 90 个水龙头。

(3) 设座问题。

甲、乙两戏院在竞争 5000 名观众，假设每个观众完全随意地选择一个戏院，且观众之间选择戏院是彼此独立的，问每个戏院至少应该设多少个座位才能保证观众因缺少座位而离开的概率小于 5%。

【解】 由于两个戏院的情况相同，故只需考虑甲戏院即可。设甲戏院需设 m 个座位，设
$$X_i=\begin{cases}1, & \text{第 } i \text{ 个观众选择甲电影院}\\ 0, & \text{其他}\end{cases},\quad i=1,2,3,\cdots,5000$$
则
$$P(X_i=1)=P(X_i=0)=0.5,i=1,2,\cdots,5000$$

若用 X 表示选择甲戏院的观众总数，则 $X=\sum_{i=1}^{5000}X_i$。

问题化为求 m 使 $P(X\geqslant m)\leqslant 0.05$，

即
$$P(X\leqslant m)\leqslant 0.95$$
因为
$$E(X_i)=\sqrt{D(X_i)}=0.5$$

由隶莫佛-拉普拉斯中心极限定理 $P(X\leqslant m)\approx\Phi\left(\dfrac{m-250}{5\sqrt{5}}\right)\geqslant 0.95$

查标准正态分布表知
$$\frac{m-250}{5\sqrt{5}}\geqslant 1.645$$
从而解得 $m\geqslant 269$，即每个戏院至少应该设 269 个座位。

(4) 抽样检验问题。

某药厂断言，该厂生产的某药品对医治一种疑难的血液病治愈率为 0.8。医院检验员任意取 100 个服用此药的病人，如果其中多于 75 个治愈，就接受这一断言，否则就拒绝这一断言。①若实际上此药对这种病的治愈率是 0.8，问接受这一断言的概率是多少？②若实际上此药对这种病的治愈率是 0.7，问接受这一断言的概率是多少？

【解】 引入随机变量
$$X_i=\begin{cases}0, & \text{第 } i \text{ 个人没有治愈}\\ 1, & \text{第 } i \text{ 个人治愈}\end{cases},\quad i=1,2,3,\cdots,100$$
X 表示抽查的 100 个人中被治愈的人数，则 $X=\sum_{i=1}^{100}X_i$。

$$① P\{X>75\}=P\left\{\sum_{i=1}^{100}X_i>75\right\}\approx P\left\{\frac{\sum X_i-100\times 0.8}{\sqrt{100\times 0.8\times 0.2}}>\frac{75-100\times 0.8}{\sqrt{100\times 0.8\times 0.2}}\right\}$$
$$\approx 1-\Phi\left(\frac{75-100\times 0.8}{\sqrt{100\times 0.8\times 0.2}}\right)=\Phi(1.25)=0.8944$$

实际治愈率为 0.8 时，接受这一断言的概率为 0.8944。

② $P\{X>75\}=P\left\{\sum_{i=1}^{100}X_i>75\right\}=P\left\{\frac{\sum X_i-100\times0.8}{\sqrt{100\times0.8\times0.2}}>\frac{75-100\times0.8}{\sqrt{100\times0.8\times0.2}}\right\}$

$$\approx 1-\Phi(1.09)=1-0.8621=0.1379$$

实际治愈率为 0.7 时，接受这一断言的概率为 0.1379。

(5) 供应问题。

假设某车间有 200 台车床独立地工作着，开工率各为 0.6，开工时耗电各为 1000W，问供电所至少要给该车间多少电力，才能以 99.9%的概率保证这个车间不会因供电不足而影响生产?

【解】 设任一时刻工作着的机床数为 X，则 X 服从参数为 $n=200$，$p=0.6$ 的二项分布，该时刻的耗电量为 XkW，如果用 k 表示供电所给该车间的最少电力，则此题所求即为 k 取何值时，有

$$P\{0\leqslant X\leqslant k\}\approx\Phi\left(\frac{k-200\times0.6}{\sqrt{200\times0.6\times0.4}}\right)-\Phi\left(\frac{0-200\times0.6}{\sqrt{200\times0.6\times0.4}}\right)=0.999$$

$$P\{0\leqslant X\leqslant k\}\approx\Phi\left(\frac{k-120}{\sqrt{48}}\right)=0.999$$

查表得$\frac{k-120}{\sqrt{48}}\approx3.1$，解得 $k=141$。

即只要给该车间 141kW 的电力，就能以 99.9%的概率保证该车间不会因电力不足而影响生产。

数学实验三

【实验目的】

学习计算随机变量的期望、方差和标准差。

【基本语句】

1. Mean[dist]

功能：求 dist 分布的数学期望。

2. Variance[dist]

功能：求 dist 分布的方差。

3. StandardDeviation[dist]

功能：求 dist 分布的标准差。

4. ExpectedValue[f,dist,x]

功能：求 dist 分布函数的数学期望 $E[f(x)]$。

【实验内容】

(1) 二项分布的期望、方差和标准差的计算。

Mathematica 语句如下：

```
Mean[BinomialDistribution[n,p]]
```

运行结果：np

Mathematica 语句如下：

```
Variance[BinomialDistribution[n,p]]
```

运行结果：np(1－p)

Mathematica 语句如下：

```
StandardDeviation[BinomialDistribution[n,p]]
```

运行结果：$\sqrt{np(1-p)}$

Mathematica 语句如下：

```
ExpectedValue[x^2,NormalDistribution[μ,σ],x]
```

运行结果：$\mu^2+\sigma^2$

Mathematica 语句如下：

```
ExpectedValue[(x-μ)^2,NormalDistribution[μ,σ],x]
```

运行结果：$3\sigma^4$

(2) 一工厂生产的某种产品的寿命(以年计)服从 $X\sim E(0.2)$的指数分布，服务承诺产品售出后一年内若损坏可以免费更换，若售出一件产品盈利 200 元，更换一个产品亏损 300 元，求工厂售出一件产品盈利的数学期望。

由题设，售出一件产品的利润 $g(x)$为

$$g(x)=\begin{cases}-300, & 0<x<1\\ 200, & x\geqslant 1\end{cases}$$

Mathematica 语句如下：

```
g = - 300UnitStep[1 - x] + 200UnitStep[x - 1];
ExpectedValue[g,ExponentialDistribution[0.2],x]
```

运行结果：109.365

即工厂平均每售出一件产品就盈利 109.365 元。

(3) 计算泊松分布的期望与方差。

Mathematica 语句如下：

```
Mean[PoissonConsulDistribution[μ,l]]
```

运行结果：$\mu/(1-\lambda)$

Mathematica 语句如下：

```
Variance[PoissonConsulDistribution[μ,λ]]
```

运行结果：$\mu/(1-\lambda)^3$

(4) 设 B 服从$[0,2\pi]$上的均匀分布，$X=\cos(B)$，$Y=\cos(A+B)$(A 为常数)，X 和 Y

的相关系数为 $\rho=\cos(A)$。产生服从 $U[0,2\pi]$ 的 N 个随机数，取 $N=100$，对应 $A=0$，$A=\frac{\pi}{3}$，$A=\frac{\pi}{2}$，$A=\pi$ 分别绘出 X 和 Y 的散点图，观察 ρ 对散点图的影响。

Mathematica 语句如下：

```
n = 100;
covar[a_] := Module[{},t1 = RandomReal[UniformDistribution[{0,2Pi}],n]];
a = 0;covar[a];
x[t_] = Cos[t];y[t_] = Cos[a + t];
txy = Table[{Cos[t1[[i]]],Cos[a + t1[[i]]]},{i,n}];
g1 = ListPlot[txy,PlotStyle RGBColor[1,0,0],DisplayFunction Identity]
a = Pi/3;covar[a];
x[t_] = Cos[t];y[t_] = Cos[a + t];
txy = Table[{Cos[t1[[i]]],Cos[a + t1[[i]]]},{i,n}];
g2 = ListPlot[txy,PlotStyle RGBColor[1,0,0],DisplayFunction Identity]
a = Pi/2;covar[a];
x[t_] = Cos[t];y[t_] = Cos[a + t];
txy = Table[{Cos[t1[[i]]],Cos[a + t1[[i]]]},{i,n}];
g3 = ListPlot[txy,PlotStyle RGBColor[1,0,0],DisplayFunction Identity]
a = Pi;covar[a];
x[t_] = Cos[t];y[t_] = Cos[a + t];
txy = Table[{Cos[t1[[i]]],Cos[a + t1[[i]]]},{i,n}];
g4 = ListPlot[txy,PlotStyle RGBColor[1,0,0],DisplayFunction Identity]
```

则依次输出 $A=0$，$A=\frac{\pi}{3}$，$A=\frac{\pi}{2}$，$A=\pi$ 时的散点图(见图 3.1)。

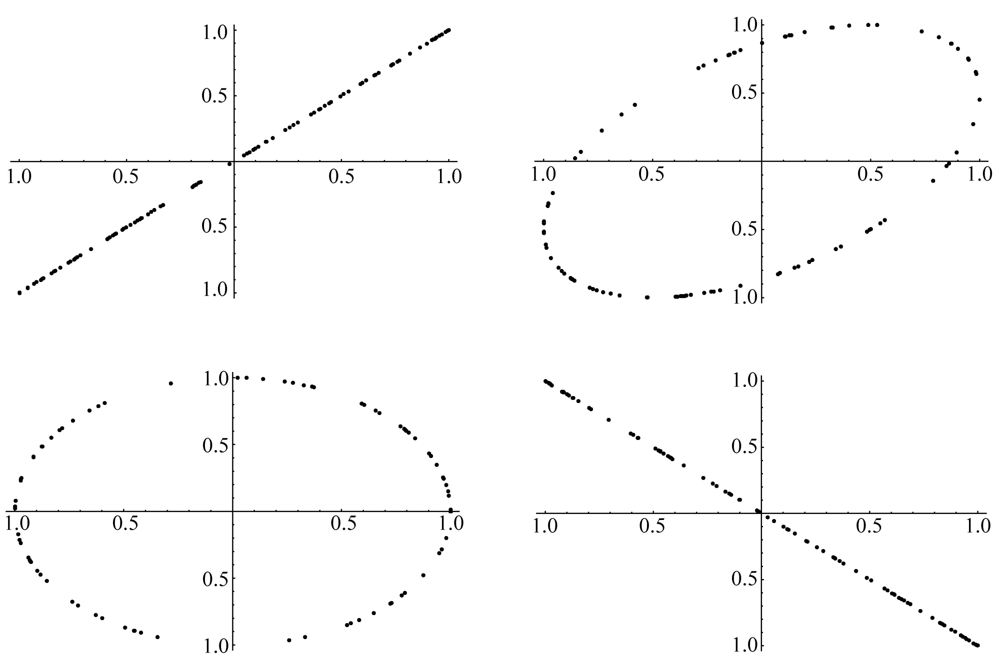


图 3.1

从上述图形可见，当$|\rho|$较大时，X和Y的线性关系较紧密，特别当$|\rho|=1$时，X和Y之间存在线性关系；当$|\rho|$较小时，X和Y的线性关系较差，特别当$\rho=0$时，X和Y不相关。

【练习】

(1) 求参数为$\lambda=5$的泊松分布的期望和方差。

(2) 求二项分布$B(10,0.3)$的期望、方差和标准差。