当 $y\leqslant 0$ 或 $y\geqslant 2$ 时，$f_Y(y)=0$，故

$$f_Y(y)=\begin{cases}1-\dfrac{y}{2}, & 0<y<2,\\ 0, & \text{其他}.\end{cases}$$

(Ⅱ) 当 $z\leqslant 0$ 时，$F_Z(z)=0$;

当 $z\geqslant 2$ 时，$F_Z(z)=1$;

当 $0<z<2$ 时，

$$F_Z(z)=P\{2X-Y\leqslant z\}=\iint\limits_{2x-y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y=z-\frac{z^2}{4}.$$

故

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

(Ⅲ) $P\left\{Y\leqslant\frac{1}{2}\,\middle|\,X\leqslant\frac{1}{2}\right\}=\dfrac{P\left\{X\leqslant\frac{1}{2},\ Y\leqslant\frac{1}{2}\right\}}{P\left\{X\leqslant\frac{1}{2}\right\}}=\dfrac{3/16}{1/4}=\dfrac{3}{4}.$

[解法 2]　(Ⅰ) 同解法 1.

(Ⅱ) 由于 $f_Z(z)=\int_{-\infty}^{+\infty}f(x,2x-z)\mathrm{d}x$，其中

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

当 $z\leqslant 0$ 或 $z\geqslant 2$ 时，$f_Z(z)=0$;

当 $0<z<2$ 时，$f_Z(z)=\int_{\frac{z}{2}}^{1}\mathrm{d}x=1-\frac{z}{2}$.

故

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

(Ⅲ) 同解法 1.

[典型错误]　在计算边缘概率密度时，弄错了积分区域. 在求 $Z=2X-Y$ 的概率密度时，误以为 X 与 Y 独立，从而用错了卷积公式.

四、随机变量的数字特征

随机变量的数字特征是用来描述随机变量分布特征的某些数字. 这些数字能从一定的角度或侧面更具体、准确而突出地刻画随机变量的性质和特点，它们包括数学期望、方差、矩、协方差和相关系数等. 本节的重点是掌握随机变量数字特征的定义、性质，并会利用数字特征的基本性质计算具体分布的数字特征. 如二项分布、几何分布、超几何分布、泊松分布、均匀分布、指数分布、一维和二维正态分布的有关数字特征. 掌握直接利用随机变量的概率分布求其函数的数学期望的方法. 掌握切比雪夫不等式，理解其意义，会用并会证明类似的不等式.

• 考试内容与要求 •

考试内容(数学三)

随机变量的数学期望(均值)、方差、标准差及其性质　随机变量函数的数学期望　切比雪夫(Chebyshev)不等式　矩、协方差、相关系数及其性质

考试要求

1. 理解随机变量数字特征(数学期望、方差、标准差、矩、协方差、相关系数)的概念，会运用数字特征的基本性质，并掌握常用分布的数字特征.

2. 会求随机变量函数的数学期望.

3. 掌握切比雪夫不等式.

考试内容(数学四)

随机变量的数学期望(均值)、方差、标准差及其性质　随机变量函数的数学期望　切比雪夫(Chebyshev)不等式　矩、协方差、相关系数及其性质

考试要求

1. 理解随机变量数字特征(数学期望、方差、标准差、矩、协方差、相关系数)的概念，会运用数字特征的基本性质，并掌握常用分布的数字特征.

2. 会求随机变量函数的数学期望.

3. 了解切比雪夫不等式.

·考试内容解析·

(一) 随机变量的数学期望

数学期望刻画了随机变量的平均值，即位置特征.

1. 定义

(1) 离散型随机变量的数学期望：设离散型随机变量 X 的概率分布为

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

如果无穷级数 $\sum\limits_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称此级数的和为随机变量 X 的数学期望，记作 $E(X)$或 EX，即

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

(2) 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 $f(x)$. 如果反常积分 $\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$ 绝对收敛,则称此积分值为随机变量 X 的数学期望,记作 EX,即

$$EX=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x.$$

如果上述级数或积分不绝对收敛，则称此随机变量的数学期望不存在.

2. 性质

(1) 对于常数 c，有 $Ec=c$.

(2) 对于常数 c 及随机变量 X，有 $E(cX)=c\cdot EX$.

(3) 设 X 与 Y 是两个随机变量. 则 $E(X\pm Y)=EX\pm EY$.

(4) 设随机变量 X 与 Y 独立，则 $E(XY)=EX\cdot EY$.

3. 随机变量函数的数学期望

设 $y=g(x)$为连续函数或分段连续函数，而 X 是任一随机变量，则随机变量 $Y=g(X)$的数学期望可以通过随机变量 X 的概率分布直接求出：

$$EY=Eg(X)=\begin{cases}\sum\limits_k g(x_k)P\{X=x_k\}, & X\text{ 是离散的},\\ \int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x, & X\text{ 是连续的}.\end{cases}$$

对于二元函数 $Z=g(X,Y)$，仍有类似公式：

$$EZ = Eg(X,Y) = \begin{cases} \sum_i \sum_j g(x_i,y_j)P\{X = x_i, Y = y_j\}, & X,Y \text{ 是离散的}, \\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y, & X,Y \text{ 是连续的}. \end{cases}$$

(二) 随机变量的方差和标准差

随机变量的方差是反映随机变量取值分散或集中程度的数字特征.

1. 定义

设 X 是一随机变量. 如果数学期望 $E(X-EX)^2$ 存在，则称之为 X 的方差，记作 DX 或 $\mathrm{Var}\,X$. 另外，$\sqrt{DX}$ 或 $\sqrt{\mathrm{Var}\,X}$ 称为 X 的标准差.

一个常用来计算方差的公式是

$$DX = EX^2 - (EX)^2 .$$

2. 性质

(1) 设 c 是常数，则 $Dc = 0$.

(2) 设 X 是随机变量，c 为常数，则

$$D(X+c) = DX, \quad D(cX) = c^2 \cdot DX .$$

(3) 设随机变量 X 与 Y 相互独立，则

$$D(X \pm Y) = DX + DY .$$

(4) $DX \geqslant 0$，并且 $DX = 0$ 当且仅当 X 以概率 1 取常数，即 $P\{X = c\} = 1$. 其中 $c = EX$.

(三) 常用随机变量的数学期望与方差

1. 如果 X 是服从 0 - 1 分布的随机变量，即 $P\{X = i\} = p^i(1-p)^{1-i}$, $i = 0,1$，则

$$EX = p, \quad DX = p(1-p).$$

2. 如果 $X \sim B(n,p)$，则

$$EX = np, \quad DX = np(1-p).$$

3. 如果 $X \sim P(\lambda)$，则

$$EX = DX = \lambda .$$

4. 如果 $X \sim U(a,b)$，则

$$EX = \frac{a+b}{2}, \quad DX = \frac{(b-a)^2}{12}.$$

5. 如果 $X \sim e(\lambda)$，则

$$EX = \frac{1}{\lambda}, \quad DX = \frac{1}{\lambda^2}.$$

6. 如果 $X \sim N(\mu,\sigma^2)$，则

$$EX = \mu, \quad DX = \sigma^2 .$$

(四) 协方差和相关系数

对于二维随机变量(X,Y)，其数字特征包括每个变量的数学期望和方差以及 X 和 Y 的联合数字特征——协方差和相关系数.

1. 定义

(1) 协方差：随机变量 X 与 Y 的协方差定义为

$$\mathrm{cov}(X,Y) = E(X-EX)(Y-EY) = E(XY) - EX \cdot EY,$$

其中

$$E(XY) = \begin{cases} \sum_i \sum_j x_i y_j P\{X = x_i, Y = y_j\}, & X,Y \text{ 是离散的}, \\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)\mathrm{d}x\mathrm{d}y, & X,Y \text{ 是连续的}. \end{cases}$$

(2) 相关系数：随机变量 X 与 Y 的相关系数定义为

$$\rho=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX\cdot DY}}=\frac{E(XY)-EX\cdot EY}{\sqrt{DX\cdot DY}}.$$

2. 协方差的性质

协方差刻画了两个随机变量间的相关性. 在相应协方差存在时. 具有如下性质:

(1) $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$.

(2) 若 X 与 Y 独立，则 $\operatorname{cov}(X,Y)=0$.

(3) 对于任意实数 a 和 b，有 $\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$.

(4) $\operatorname{cov}(X+Y,Z)=\operatorname{cov}(X,Z)+\operatorname{cov}(Y,Z)$.

(5) $|\operatorname{cov}(X,Y)|\leqslant\sqrt{DX}\sqrt{DY}$.

(6) $D(X\pm Y)=DX+DY\pm 2\operatorname{cov}(X,Y)$.

3. 相关系数的性质

相关系数刻画了两个随机变量间的线性相关性，它有如下性质：

(1) $-1\leqslant\rho\leqslant 1$.

(2) 如果 X 与 Y 独立. 则 $\rho=0$.

(3) $|\rho|=1$ 的充要条件是 X 和 Y 以概率 1 互为线性函数.

4. 随机变量的相关性

设两个随机变量 X 与 Y 间的相关系数 ρ 存在. 如果 $\rho=0$，则称 X 与 Y 不相关，否则就称 X 与 Y 相关. 如果 $\rho>0$. 则称 X 与 Y 正相关，否则称 X 与 Y 负相关.

两独立随机变量一定不相关，但是两个不相关的随机变量却未必独立. 然而，对于联合分布是二维正态分布的随机变量. 其独立性与不相关性是等价的.

(五) 矩

对于随机变量 X，若

$$EX^k,\quad k=1,2,\cdots$$

存在，则称之为 X 的 k 阶原点矩. 如果

$$E(X-EX)^k,\quad k=1,2,\cdots$$

存在. 则称之为 X 的 k 阶中心矩.

原点矩和中心矩可以互为表出. 方差即为二阶中心矩.

(六) 切比雪夫不等式

设随机变量 X 的数学期望 EX 和方差 DX 都存在. 则对于任意的 $\varepsilon>0$. 有

$$P\{|X-EX|\geqslant\varepsilon\}\leqslant\frac{DX}{\varepsilon^2}.$$

·例题详解·

例 3.4.1 设随机变量 X 服从参数为 λ 的指数分布，则 $P\{X>\sqrt{DX}\}=$ ________.

[答案] $\frac{1}{e}$.

[提示] 本题主要考查常见随机变量的概率分布和数字特征. 首先利用数字特征的求解方法，可以求得 $DX=\frac{1}{\lambda^2}$. 之后再利用概率的计算方法可以求得此概率.

[解] 由 X 服从参数为 λ 的指数分布，知

$$DX=\frac{1}{\lambda^2}.$$

则

$$P\{X>\sqrt{DX}\}=P\left\{X>\frac{1}{\lambda}\right\}=\int_{\frac{1}{\lambda}}^{+\infty}\lambda e^{-\lambda x}dx=\frac{1}{e}.$$

[典型错误] 没能理解 DX 只是一个特殊常数，从而不知如何求解题中的概率；也有部分考生没能正确记住指数分布的概率密度和数字特征，从而得到错误的答案.

例 3.4.2 设随机变量 X 和 Y 的联合概率分布为：

X \ 概率 \ Y	-1	0	1
0	0.07	0.18	0.15
1	0.08	0.32	0.20

则 X^2 和 Y^2 的协方差 $\operatorname{cov}(X^2,Y^2)=$________.

[答案] -0.02.

[提示] 本题主要考查：①会利用离散型二维随机变量(X,Y)的联合概率分布求 X^2，Y^2 及 X^2Y^2 等 X，Y 的简单函数的概率分布. ②会根据数学期望的定义求相关随机变量的数学期望. ③知道 $\operatorname{cov}(X^2,Y^2)=E(X^2Y^2)-EX^2\cdot EY^2$. 如果考生直接利用协方差公式 $\operatorname{cov}(X^2,Y^2)=E(X^2-EX^2)(Y^2-EY^2)$计算，会复杂得多.

[解] 根据随机变量 X 和 Y 的联合概率分布，则 X^2，Y^2 和 X^2Y^2 的概率分布分别为：

$$X^2\sim\begin{pmatrix}0 & 1\\ 0.4 & 0.6\end{pmatrix},$$

$$Y^2\sim\begin{pmatrix}0 & 1\\ 0.5 & 0.5\end{pmatrix},$$

$$X^2Y^2\sim\begin{pmatrix}0 & 1\\ 0.72 & 0.28\end{pmatrix}.$$

故
$$EX^2=0.6,\ EY^2=0.5,\ E(X^2Y^2)=0.28,$$

因而
$$\operatorname{cov}(X^2,Y^2)=E(X^2Y^2)-EX^2\cdot EY^2=0.28-0.6\times0.5=-0.02.$$

[典型错误] 不知如何计算 X^2，Y^2 的边缘分布及 X^2Y^2 的概率分布.

例 3.4.3 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________.

[答案] 0.9.

[提示] 本题主要考查两个随机变量间相关系数的概念及其运算，随机变量的数学期望及方差的计算. 如果对相关系数的概念比较清楚，此题的答案不用计算即可知仍为 0.9.

[解] 由于
$$\rho_{YZ}=\frac{E[(Y-EY)(Z-EZ)]}{\sqrt{DY\cdot DZ}},$$

又
$$DZ=D(X-0.4)=DX,$$
$$Z-EZ=(X-0.4)-E(X-0.4)=X-EX,$$

所以
$$\rho_{YZ}=\frac{E[(Y-EY)(X-EX)]}{\sqrt{DY\cdot DX}}=\rho_{YX}=\rho_{XY}=0.9.$$

[典型错误] 由 X 与 Y 的相关系数为 0.9，而误以为 X 与 $Z=X-0.4$ 的相关系数为 $0.9-0.4=0.5$.

例 3.4.4 设随机变量 X_{ij} $(i,j=1,2,\cdots,n;n\geqslant2)$独立同分布，$EX_{ij}=2$，则行列式

$$Y=\begin{vmatrix}X_{11} & X_{12} & \cdots & X_{1n}\\ X_{21} & X_{22} & \cdots & X_{2n}\\ \vdots & \vdots & & \vdots\\ X_{n1} & X_{n2} & \cdots & X_{nn}\end{vmatrix}$$

的数学期望 $EY=$________.

[答案] 0.

[提示] 本题主要考查行列式的计算、随机变量的独立性及期望．只要把 Y 展成 X_{ij} 的乘积和的形式，则由独立性及数学期望的性质不难求得 $EY=0$.

[解] 由于 $Y=\sum(-1)^{\tau(i_1i_2\cdots i_n)}X_{1i_1}X_{2i_2}\cdots X_{ni_n}$，其中 τ 为 $i_1i_2\cdots i_n$ 的逆序数，且随机变量 X_{ij} $(i,j=1,2,\cdots,n)$ 独立同分布，因此

$$EY=\sum(-1)^{\tau}EX_{1i_1}\cdot EX_{2i_2}\cdot\cdots\cdot EX_{ni_n}$$

$$=\begin{vmatrix} EX_{11} & EX_{12} & \cdots & EX_{1n} \\ EX_{21} & EX_{22} & \cdots & EX_{2n} \\ \vdots & \vdots & & \vdots \\ EX_{n1} & EX_{n2} & \cdots & EX_{nn}\end{vmatrix}=\begin{vmatrix} 2 & 2 & \cdots & 2 \\ 2 & 2 & \cdots & 2 \\ \vdots & \vdots & & \vdots \\ 2 & 2 & \cdots & 2\end{vmatrix}=0.$$

[典型错误] 不会计算 Y 的行列式.

例 3.4.5 设随机变量 X 服从参数为 λ 的泊松分布，且已知 $E[(X-1)(X-2)]=1$，则 $\lambda=$ ________.

[答案] 1.

[提示] 本题主要考查泊松分布及其数字特征，只要打开 $(X-1)(X-2)=X^2-3X+2$，之后再利用数学期望的性质即可求得.

[解] 由于 $X\sim P(\lambda)$，则 $EX=\lambda$，$DX=\lambda$．而

$$\begin{aligned}E[(X-1)(X-2)]&=EX^2-3EX+2\\&=DX+(EX)^2-3EX+2\\&=\lambda+\lambda^2-3\lambda+2=1,\end{aligned}$$

解得 $\lambda=1$.

[典型错误] 不会利用数学期望的性质进行计算.

例 3.4.6 设随机变量 X 在区间 $[-1,2]$ 上服从均匀分布，随机变量

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

则方差 $DY=$ ________.

[答案] $\dfrac{8}{9}$.

[提示] 本题主要考查均匀分布、随机变量的期望与方差．由 Y 的定义，易得

$$EY=1\cdot P\{X>0\}+0\cdot P\{X=0\}+(-1)\cdot P\{X<0\},$$

之后再利用定义求得 DY.

[解] 随机变量 X 在区间 $[-1,2]$ 上服从均匀分布，则

$$f(x)=\begin{cases}\dfrac{1}{3}, & x\in[-1,2],\\ 0, & x\notin[-1,2].\end{cases}$$

因此

$$EY=1\cdot P\{x>0\}+0\cdot P\{x=0\}-P\{x<0\}=\frac{2}{3}-\frac{1}{3}=\frac{1}{3},$$

$$\begin{aligned}DY&=(1-EY)^2\cdot P\{x>0\}+(0-EY)^2\cdot P\{x=0\}+(-1-EY)^2\cdot P\{x<0\}\\&=\frac{4}{9}\cdot\frac{2}{3}+\frac{1}{9}\times0+\frac{16}{9}\times\frac{1}{3}=\frac{8}{9}.\end{aligned}$$

[典型错误] 不会计算分段函数的期望与方差.

例 3.4.7 设随机变量 X 和 Y 的相关系数为 0.5，$EX=EY=0$，$EX^2=EY^2=2$．则 $E(X+Y)^2=$ ________.

[答案] 6.

[提示] 本题主要考查随机变量数字特征(期望、方差、协方差和相关系数)的基本性质．只要展开 $E(X+Y)^2=EX^2+2E(XY)+EY^2$，之后再由相关系数求得 $E(XY)$ 即可.

[解] $E(X+Y)^2=EX^2+2E(XY)+EY^2$. 显然只需求得 $E(XY)$的值.

由相关系数 $\rho=0.5$，得

$$\rho=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX}\cdot\sqrt{DY}}=\frac{E(XY)-EX\cdot EY}{\sqrt{DX}\cdot\sqrt{DY}}=\frac{E(XY)}{\sqrt{DX}\cdot\sqrt{DY}}=0.5,$$

所以 $E(XY)=0.5\sqrt{DX}\cdot\sqrt{DY}$. 又

$$DX=EX^2-(EX)^2=EX^2=2,$$
$$DY=EY^2-(EY)^2=EY^2=2,$$

故 $E(XY)=0.5\cdot\sqrt{2}\cdot\sqrt{2}=1$，于是

$$E(X+Y)^2=EX^2+2E(XY)+EY^2=2+2+2=6.$$

[典型错误] 不会利用已知条件计算$E(XY)$.

例 3.4.8 设随机变量$X_1,X_2,\cdots,X_n$ $(n>1)$独立同分布. 且其方差为 $\sigma^2>0$. 令 $Y=\frac{1}{n}\sum_{i=1}^{n}X_i$，则(　　).

(A) $\operatorname{cov}(X_1,Y)=\frac{\sigma^2}{n}$　　(B) $\operatorname{cov}(X_1,Y)=\sigma^2$

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

[答案] (A).

[提示] 本题主要考查随机变量的方差、协方差的计算及独立性. 要注意随机变量和的方差不一定等于各自方差的和.

[解]

$$\begin{aligned}\operatorname{cov}(X_1,Y)&=\operatorname{cov}\left(X_1,\frac{1}{n}\sum_{i=1}^{n}X_i\right)\\&=\frac{1}{n}\operatorname{cov}(X_1,X_1)+\frac{1}{n}\sum_{i=2}^{n}\operatorname{cov}(X_1,X_i)\\&=\frac{1}{n}DX_1=\frac{\sigma^2}{n}.\end{aligned}$$

此处用到：因为 X_1 与 X_i $(i=2,3,\cdots,n)$独立，所以

$$\operatorname{cov}(X_1,X_i)=E(X_1X_i)-EX_1\cdot EX_i=0.$$

故应选(A).

注意：

$$\begin{aligned}D(X_1+Y)&=D\left(\frac{n+1}{n}X_1+\frac{1}{n}X_2+\cdots+\frac{1}{n}X_n\right)\\&=\frac{(n+1)^2}{n^2}\sigma^2+\frac{n-1}{n^2}\sigma^2=\frac{n+3}{n}\sigma^2,\\D(X_1-Y)&=D\left(\frac{n-1}{n}X_1-\frac{1}{n}X_2-\cdots-\frac{1}{n}X_n\right)\\&=\frac{(n-1)^2}{n^2}\sigma^2+\frac{n-1}{n^2}\sigma^2=\frac{n-1}{n}\sigma^2,\end{aligned}$$

显然不能选(C)或(D).

[典型错误] 选(D)，误以为 X_1 与 Y 独立. 所以

$$D(X_1-Y)=DX_1+DY=\sigma^2+\frac{1}{n}\sigma^2=\frac{n+1}{n}\sigma^2.$$

例 3.4.9 将一枚硬币重复掷 n 次，以 X 和 Y 分别表示正面向上和反面向上的次数，则 X 和 Y 的相关系数等于(　　).

(A) -1　　(B) 0　　(C) $\frac{1}{2}$　　(D) 1

[答案] (A).

[提示] 本题主要考查 n 重伯努利试验、相关系数的计算. 解题的关键在于明确 X 与 Y 间的关系：$X+Y=n$，即 $Y=n-X$. 由此关系可知 Y 与 X 是负相关的，故只有选项(A)是正确的.

[解]
$$DY=D(n-X)=DX,$$
$$\operatorname{cov}(X,Y)=\operatorname{cov}(X,n-X)=-\operatorname{cov}(X,X)=-DX.$$
因此
$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX}\cdot\sqrt{DY}}=\frac{-DX}{DX}=-1.$$

[典型错误] 不清楚 X 与 Y 之间的关系：$X+Y=n$.

例 3.4.10 设两个相互独立的随机变量 X 和 Y 的方差分别为 4 和 2. 则随机变量 $3X-2Y$ 的方差是(　　).

(A) 8　　(B) 16　　(C) 28　　(D) 44

[答案] (D).

[提示] 本题主要考查随机变量函数的数字特征. 注意到 X 与 Y 是独立的，则知道 $D(3X-2Y)=9\cdot DX+4\cdot DY$.

[解] 由于 X 与 Y 是独立的，故
$$D(3X-2Y)=3^2\cdot DX+2^2\cdot DY=3^2\cdot 4+2^2\cdot 2=44.$$

[典型错误] 上述(A)、(B)、(C)三个干扰项的设置分别是针对考生平时易犯的三种错误倾向的，即误以为 $D(3X-2Y)$等于 $3DX-2DY$，$3DX+2DY$ 和 3^2DX-2^2DY.

例 3.4.11 设 X 是一随机变量. $EX=\mu$. $DX=\sigma^2$ (μ,σ 为大于 0 的常数). 则对任意常数 c. 必有(　　).

(A) $E(X-c)^2=EX^2-c^2$　　(B) $E(X-c)^2=E(X-\mu)^2$

(C) $E(X-c)^2<E(X-\mu)^2$　　(D) $E(X-c)^2\geqslant E(X-\mu)^2$

[答案] (D).

[提示] 本题主要考查随机变量方差的性质. 如果熟悉方差的性质，易知选项(D)正确.

[解]
$$\begin{aligned}E(X-c)^2&=E(X-\mu+\mu-c)^2\\&=E(X-\mu)^2+E(\mu-c)^2+2(\mu-c)E(X-\mu)\\&=E(X-\mu)^2+(\mu-c)^2\\&\geqslant E(X-\mu)^2.\end{aligned}$$

[典型错误] 弄错了不等号的方向而选(C).

例 3.4.12 设随机变量 X 和 Y 的方差存在且不等于 0，则 $D(X+Y)=DX+DY$ 是 X 和 Y (　　).

(A) 不相关的充分条件. 但不是必要条件　　(B) 独立的充分条件，但不是必要条件

(C) 不相关的充分必要条件　　(D) 独立的充分必要条件

[答案] (C).

[提示] 本题主要考查随机变量的数字特征、不相关与独立的概念及关系. 由已知条件 $D(X+Y)=DX+DY$ 仅可以推出二者的相关系数为零，而不能推出二者独立.

[解]
$$\begin{aligned}D(X+Y)&=E(X+Y-EX-EY)^2\\&=E(X-EX)^2+E(Y-EY)^2+2E(X-EX)(Y-EY)\\&=DX+DY+2\sqrt{DX}\cdot\sqrt{DY}\cdot\rho_{XY}.\end{aligned}$$
由 $D(X+Y)=DX+DY$ 知 $\rho_{XY}=0$.

[典型错误] 混淆独立与不相关这两个概念.

例 3.4.13 设二维随机变量(X,Y)服从二维正态分布，则随机变量 $\xi=X+Y$ 与 $\eta=X-Y$ 不相关的充分必要条件为(　　).

(A) $EX=EY$

(B) $EX^2-(EX)^2=EY^2-(EY)^2$

(C) $EX^2=EY^2$

(D) $EX^2+(EX)^2=EY^2+(EY)^2$

[答案] (B).

[提示] 本题主要考查随机变量的函数、随机变量相关与不相关的性质. 只要验证与 $\operatorname{cov}(\xi,\eta)=0$ 等价的条件即可选择正确的答案.

[解] $\operatorname{cov}(\xi,\eta)=E(X+Y)(X-Y)-E(X+Y)\cdot E(X-Y)$

$$=EX^2-EY^2-(EX)^2+(EY)^2.$$

因为 ξ, η 不相关$\Leftrightarrow\operatorname{cov}(\xi,\eta)=0\Leftrightarrow EX^2-(EX)^2=EY^2-(EY)^2$, 故选项(B)正确.

[典型错误] 从(X,Y)服从二维正态分布出发研究 ξ 与 η 的相关性. 实际上, 本题所给的二维正态分布这一条件是多余的.

例 3.4.14 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: (Ⅰ) 二维随机变量(X,Y)的概率分布;

(Ⅱ) X 与 Y 的相关系数 ρ_{XY};

(Ⅲ) $Z=X^2+Y^2$ 的概率分布.

[提示] 本题主要考查概率的运算公式、二维离散型随机变量的概率分布及其相关系数的求解、二维离散型随机变量简单函数的概率分布的求解.

[解] (Ⅰ) $$P(AB)=P(A)P(B|A)=\frac{1}{12},\ P(B)=\frac{P(AB)}{P(A|B)}=\frac{1}{6},$$

则 $$P\{X=1,Y=1\}=P(AB)=\frac{1}{12},$$

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

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

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

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

$\left(\text{或 } P\{X=0,Y=0\}=1-\frac{1}{12}-\frac{1}{6}-\frac{1}{12}=\frac{2}{3}\right)$,

即(X,Y)的概率分布为

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

(Ⅱ) [解法 1] $$EX=P(A)=\frac{1}{4},\ EY=P(B)=\frac{1}{6},\ E(XY)=\frac{1}{12},$$

则 $$\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY=\frac{1}{24},$$

$$EX^2=P(A)=\frac{1}{4},\ EY^2=P(B)=\frac{1}{6},$$

$$DX=EX^2-(EX)^2=\frac{3}{16},\ DY=EY^2-(EY)^2=\frac{5}{36},$$

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX\cdot DY}}=\frac{1}{\sqrt{15}}.$$

[解法 2] 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 $$EX=\frac{1}{4},\ EY=\frac{1}{6},\ 而\ E(XY)=\frac{1}{12}.$$

故 $$\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY=\frac{1}{24}.$$

$$EX^2=\frac{1}{4},\qquad EY^2=\frac{1}{6},$$

$$DX=EX^2-(EX)^2=\frac{3}{16},\ DY=EY^2-(EY^2)=\frac{5}{36}.$$

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX\cdot DY}}=\frac{1}{\sqrt{15}}.$$

(Ⅲ) Z 的可能取值为 0，1，2，则有

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

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

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

即 Z 的概率分布为

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

[典型错误]

① 不会把求(X,Y)的分布转化成计算随机事件的概率.

② 对计算二维离散型随机变量的相关系数不熟练.

③ 由于运算量较大，出现不应有的计算错误.

④ 未能把 Z 正确理解为一维离散型随机变量，因为常见的题型多是求 Z 的概率密度.

例 3.4.15 设随机变量 X 的概率密度为

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

对 X 独立地重复观察 4 次，用 Y 表示观察值大于$\frac{\pi}{3}$的次数，求 Y^2 的数学期望.

[提示] 本题主要考查连续型随机变量的分布函数、独立重复试验的数学期望和方差.

[解法 1] $$P\left\{X>\frac{\pi}{3}\right\}=1-P\left\{X\leqslant\frac{\pi}{3}\right\}=1-\int_{-\infty}^{\frac{\pi}{3}}f(x)\mathrm{d}x$$

$$=1-\int_0^{\frac{\pi}{3}}\frac{1}{2}\cos\frac{x}{2}\mathrm{d}x=1-\sin\frac{x}{2}\Big|_0^{\frac{\pi}{3}}$$

$$=1-\sin\frac{\pi}{6}=\frac{1}{2},$$

所以 $Y\sim B\left(4,\frac{1}{2}\right)$，从而

$$EY=np=4\cdot\frac{1}{2}=2,$$

$$DY=npq=4\cdot\frac{1}{2}\left(1-\frac{1}{2}\right)=1,$$

所以
$$EY^2=DY+(EY)^2=1+2^2=5.$$

［解法 2］ $P\left\{X>\frac{\pi}{3}\right\}=\frac{1}{2}$同上，$Y$ 的分布律如下：

Y	0	1	2	3	4
P	$\left(1-\frac{1}{2}\right)^4$	$C_4^1\left(\frac{1}{2}\right)\left(1-\frac{1}{2}\right)^3$	$C_4^2\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right)^2$	$C_4^3\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)$	$C_4^4\left(\frac{1}{2}\right)^4$

所以
$$EY^2=0+C_4^1\left(\frac{1}{16}\right)+C_4^2\left(\frac{1}{16}\right)\cdot 4+C_4^3\left(\frac{1}{16}\right)\cdot 9+C_4^4\left(\frac{1}{16}\right)\cdot 16=5.$$

［典型错误］

① 不会计算$P\left\{X>\frac{\pi}{3}\right\}$.

② 甚至积分$\int_0^{\frac{\pi}{3}}\frac{1}{2}\cos\frac{x}{2}\mathrm{d}x$ 也计算错.

③ 不知道 $Y\sim B\left(4,\frac{1}{2}\right)$或者不会进一步去套用公式 $EY=np$、$DY=np(1-p)$.

④ 不知道③中所说，但也不会直接写出 Y 的分布律计算 EY^2.

例 3.4.16 假设随机变量 U 在区间$[-2,2]$上服从均匀分布. 随机变量

$$X=\begin{cases}-1, & 若\ U\leqslant -1,\\ 1, & 若\ U>-1;\end{cases}\quad Y=\begin{cases}-1, & 若\ U\leqslant 1,\\ 1, & 若\ U>1.\end{cases}$$

试求：（Ⅰ）X 和 Y 的联合概率分布；

（Ⅱ）$D(X+Y)$.

［提示］ 本题主要考查随机变量的联合分布、随机变量函数的数字特征. 首先应通过所给的 X，Y 的取值，判断出(X,Y)的几种可能值以及取这些可能值的条件；其次熟悉均匀分布的定义，并会求在任一点 X 处的分布值$P\{U<X\}$. 最后，会通过(X,Y)的联合分布写出 $X+Y$ 和$(X+Y)^2$ 的分布，从而利用 $D(X+Y)=E(X+Y)^2-[E(X+Y)]^2$ 求出 $D(X+Y)$.

［解］（Ⅰ）二维随机变量(X,Y)有四个可能值：$(-1,-1)$，$(-1,1)$，$(1,-1)$，$(1,1)$.

$$P\{X=-1,Y=-1\}=P\{U\leqslant -1,U\leqslant 1\}=\frac{1}{4},$$

$$P\{X=-1,Y=1\}=P\{U\leqslant -1,U>1\}=0,$$

$$P\{X=1,Y=-1\}=P\{U>-1,U\leqslant 1\}=\frac{1}{2},$$

$$P\{X=1,Y=1\}=P\{U>-1,U>1\}=\frac{1}{4}.$$

于是，得 X 和 Y 的联合概率分布为

$$(X,Y)\sim\begin{pmatrix}(-1,-1) & (-1,1) & (1,-1) & (1,1)\\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4}\end{pmatrix}.$$

（Ⅱ）$X+Y$ 和$(X+Y)^2$ 的概率分布相应为

$$X+Y\sim\begin{pmatrix}-2 & 0 & 2\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{pmatrix},\quad (X+Y)^2\sim\begin{pmatrix}0 & 4\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix},$$

由此可见

$$E(X+Y)=-\frac{2}{4}+\frac{2}{4}=0,$$

$$D(X+Y)=E(X+Y)^2=2.$$

[典型错误] 如果考生平时对这类题目缺乏练习的话，就会被所给的多个随机变量之间的关系搞混淆，从而无从下手．特别是对离散型随机变量的处理在平时教学中或练习中也常被忽视．

例 3.4.17 已知甲、乙两箱中装有同种产品，其中甲箱中装有 3 件合格品和 3 件次品，乙箱中仅装有 3 件合格品．从甲箱中任取 3 件产品放入乙箱后，求：

(Ⅰ) 乙箱中次品件数 X 的数学期望；

(Ⅱ) 从乙箱中任取一件产品是次品的概率．

[提示] 本题主要考查古典型概率的计算、离散型随机变量的数学期望、全概率公式．由古典型概率的计算方法可以求得 X 的概率分布，之后再利用全概率公式求解第(Ⅱ)问．

[解] (Ⅰ) 乙箱中次品件数 X 是个随机变量，X 的取值为 0，1，2，3．X 的概率分布为

$$P\{X=k\}=\frac{C_3^kC_3^{3-k}}{C_6^3},$$

列表得

X	0	1	2	3
p	$\frac{1}{20}$	$\frac{9}{20}$	$\frac{9}{20}$	$\frac{1}{20}$

因此

$$EX=0\times\frac{1}{20}+1\times\frac{9}{20}+2\times\frac{9}{20}+3\times\frac{1}{20}=\frac{3}{2}.$$

(Ⅱ) 设 A 表示事件“从乙箱中任取一件产品是次品”．由全概率公式，有

$$P(A)=\sum_{k=0}^{3}P\{X=k\}\cdot P\{A|X=k\}$$

$$=\frac{1}{20}\cdot 0+\frac{9}{20}\cdot\frac{1}{6}+\frac{9}{20}\cdot\frac{2}{6}+\frac{1}{20}\cdot\frac{3}{6}=\frac{1}{4}.$$

[典型错误]

① 错误地认为 $X\sim B\left(3,\frac{1}{2}\right)$．随机变量 $X\sim B(n,p)$是指：每次成功的概率为 p，失败的概率为$1-p$，如此独立地重复试验 n 次，成功的次数 $X\sim B(n,p)$．而现在从甲箱中任取 3 件放入乙箱中，指的显然是不放回抽取．

② 全概率公式写错．

③ 数字计算错．

④ 缺少概率描述而猜出答案．例如：(a)从甲袋中摸出的每一个球是次品的概率为$\frac{1}{2}$，共摸了 3 个球，平均有 $3\times\frac{1}{2}$个次品；(b)乙袋中平均有$\frac{3}{2}$个次品，共 6 个球．所以从乙袋中摸到 1 个次品的概率为$\frac{1.5}{6}=\frac{1}{4}$．以上(a)、(b)缺乏概率理论与计算．实际上，考生将数学期望的常识性理解用来作为论据．

例 3.4.18 设二维随机变量(X,Y)的密度函数为

$$f(x,y)=\frac{1}{2}[\varphi_1(x,y)+\varphi_2(x,y)],$$

其中 $\varphi_1(x,y)$和 $\varphi_2(x,y)$都是二维正态密度函数，且它们对应的二维随机变量的相关系数分别为$\frac{1}{3}$和$-\frac{1}{3}$，它们的边缘密度函数所对应的随机变量的数学期望都是零，方差都是 1．

(Ⅰ) 求随机变量 X 和 Y 的密度函数$f_1(x)$和 $f_2(y)$及 X 和 Y 的相关系数ρ(可以直接利用二维正态密度的性质)；

(Ⅱ) 问 X 和 Y 是否独立？为什么？

[提示] 本题主要考查如何根据二维随机变量(X,Y)的密度函数求 X, Y 的边缘密度；二维正态分布密度与两个随机变量的均值、方差、相关系数之间的关系；随机变量独立性与相关系数之间的关系；随机变量之间的独立性的判别准则等知识点. 首先应正确理解题意. 要清楚密度函数 $\varphi_1(x,y)$, $\varphi_2(x,y)$和 $f(x,y)$分别是三个不同的二维随机变量(X_1,Y_1), (X_2,Y_2)及(X,Y)的密度. 由题意知(X_1,Y_1), (X_2,Y_2)都服从正态分布, 且 $EX_1=EX_2=0$, $DX_1=DX_2=1$, $EY_1=EY_2=0$, $DY_1=DY_2=1$, 根据 X_1 与 Y_1 的相关系数$\left(\rho_1=\frac{1}{3}\right)$、$X_2$ 与 Y_2 的相关系数$\left(\rho_2=-\frac{1}{3}\right)$可以写出$\varphi_1(x,y)$和 $\varphi_2(x,y)$的表达式. X 和 Y 是否独立, 视 $f(x,y)$是否等于 $f_1(x)f_2(y)$而定, 而不能由 ρ 是否为零来确定. 因为$f(x,y)$并非是二维正态分布密度函数.

[解] (Ⅰ) 由于二维正态密度函数的两个边缘密度都是正态密度函数, 因此 $\varphi_1(x,y)$和 $\varphi_2(x,y)$的两个边缘密度为标准正态密度函数, 故

$$f_1(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\frac{1}{2}\left[\int_{-\infty}^{+\infty}\varphi_1(x,y)\mathrm{d}y+\int_{-\infty}^{+\infty}\varphi_2(x,y)\mathrm{d}y\right]$$

$$=\frac{1}{2}\left[\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}+\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\right]=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}.$$

同理, 得

$$f_2(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}.$$

由于 $X\sim N(0,1)$, $Y\sim N(0,1)$, 可见 $EX=EY=0$, $DX=DY=1$. 随机变量 X 和 Y 的相关系数

$$\begin{aligned}\rho&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\,\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2}\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\varphi_1(x,y)\mathrm{d}x\mathrm{d}y+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\varphi_2(x,y)\,\mathrm{d}x\mathrm{d}y\right]\\&=\frac{1}{2}\left(\frac{1}{3}-\frac{1}{3}\right)=0.\end{aligned}$$

(Ⅱ) 由题设

$$f(x,y)=\frac{3}{8\pi\sqrt{2}}\left[\mathrm{e}^{-\frac{9}{16}\left(x^2-\frac{2}{3}xy+y^2\right)}+\mathrm{e}^{-\frac{9}{16}\left(x^2+\frac{2}{3}xy+y^2\right)}\right],$$

$$f_1(x)\cdot f_2(y)=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2}{2}}\cdot\mathrm{e}^{-\frac{y^2}{2}}=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}.$$

$$f(x,y)\neq f_1(x)\cdot f_2(y),$$

所以 X 与 Y 不独立.

[典型错误] 题意理解不透, 不知如何下手.

例 3.4.19 从学校乘汽车到火车站的途中有 3 个交通岗, 假设在各个交通岗遇到红灯的事件是相互独立的, 并且概率都是$\frac{2}{5}$. 设 X 为途中遇到红灯的次数, 求随机变量 X 的分布律、分布函数和数学期望.

[提示] 本题主要考查二项分布、离散型随机变量的分布函数及数学期望, 是二项分布在实际中应用的简单实例. 因为 $X\sim B\left(3,\frac{2}{5}\right)$, 则

$$P\{X=k\}=\mathrm{C}_3^k\left(\frac{2}{5}\right)^k\left(\frac{3}{5}\right)^{3-k},\ k=0,1,2,3.$$

对于离散型的随机变量, 一般不提密度函数而提分布律. 即 X 在不同点取值的概率. 在求分布函数时注意

$$F(x)=P\{X\leqslant x\}=\sum_{\{k\mid k\leqslant x\}}P\{X=k\}.$$

若记住 $EX=np$ 这一公式, 求数学期望就很简单了.

[解] X 服从二项分布 $B\left(3,\frac{2}{5}\right)$，$X$ 的可能取值为 0,1,2,3．从而

$$P\{X=0\}=C_3^0\left(1-\frac{2}{5}\right)^3=\frac{27}{125},$$

$$P\{X=1\}=C_3^1\cdot\frac{2}{5}\cdot\left(1-\frac{2}{5}\right)^2=\frac{54}{125},$$

$$P\{X=2\}=C_3^2\left(\frac{2}{5}\right)^2\left(1-\frac{2}{5}\right)=\frac{36}{125},$$

$$P\{X=3\}=C_3^3\left(\frac{2}{5}\right)^3=\frac{8}{125},$$

即 X 的分布律为

X	0	1	2	3
P	$\frac{27}{125}$	$\frac{54}{125}$	$\frac{36}{125}$	$\frac{8}{125}$

因此，X 的分布函数为

$$F(x)=P\{X\leqslant x\}=\begin{cases}0, & x<0,\\ \frac{27}{125}, & 0\leqslant x<1,\\ \frac{81}{125}, & 1\leqslant x<2,\\ \frac{117}{125}, & 2\leqslant x<3,\\ 1, & x\geqslant 3.\end{cases}$$

X 的数学期望为 $EX=3\cdot\frac{2}{5}=\frac{6}{5}\left(\text{或 } EX=0\cdot\frac{27}{125}+1\cdot\frac{54}{125}+2\cdot\frac{36}{125}+3\cdot\frac{8}{125}=\frac{6}{5}\right)$.

[典型错误]

① 分布函数的概念不清楚，从而写不对，主要问题是将分布函数写成分段函数时，写错分段区间.

② 计算错误.

例 3.4.20 游客乘电梯从底层到电视塔顶层观光，电梯于每个整点的第 5、25 和 55 分钟从底层起行．假设一游客在早八点的第 X 分钟到达底层候梯处，且 X 在[0,60]上均匀分布，求该游客等候时间的数学期望.

[提示] 本题主要考查随机变量函数的数学期望．通过认真分析题意知，游客等候时间是他到达候梯处时间 X 的函数，这一函数是个分段函数．列出这一函数后，求游客等候时间的数学期望就转化成求简单随机变量函数的数学期望问题．这里要求考生会写出[0,60]上均匀分布的密度函数表示式.

[解] 已知 X 在[0,60]上服从均匀分布，其密度函数为

$$f(x)=\begin{cases}\frac{1}{60}, & \text{若 } 0\leqslant x\leqslant 60,\\ 0, & \text{其他}.\end{cases}$$

设 Y 是游客等候电梯的时间(单位:分)，则

$$Y=g(X)=\begin{cases}5-X, & \text{若 } 0<X\leqslant 5,\\ 25-X, & \text{若 } 5<X\leqslant 25,\\ 55-X, & \text{若 } 25<X\leqslant 55,\\ 60-X+5, & \text{若 } 55<X\leqslant 60,\end{cases}$$

因此

$$\begin{aligned}EY&=Eg(X)\\&=\int_{-\infty}^{+\infty}g(x)\cdot f(x)\mathrm{d}x=\frac{1}{60}\int_0^{60}g(x)\mathrm{d}x\end{aligned}$$

$$= \frac{1}{60}\left[\int_0^5 (5-x)\mathrm{d}x + \int_5^{25}(25-x)\mathrm{d}x + \int_{25}^{55}(55-x)\mathrm{d}x + \int_{55}^{60}(65-x)\mathrm{d}x\right]$$
$$= \frac{1}{60}(12.5+200+450+37.5)$$
$$\approx 11.67.$$

[典型错误] 不知如何用 X 表示等候时间这一随机变量.

例 3.4.21 两台同样的自动记录仪，每台无故障工作的时间服从参数为 5 的指数分布. 首先开动其中一台，当其发生故障时停用而另一台自行开动. 试求两台记录仪无故障工作的总时间 T 的概率密度 $f(t)$、数学期望和方差.

[提示] 本题主要考查指数分布、两独立随机变量和的分布、卷积公式、独立随机变量和的期望与方差. 根据题意，总时间 T 是第一台记录仪无故障工作时间 X_1 和第二台记录仪无故障工作时间 X_2 的和，即 $T=X_1+X_2$. 且 X_1，X_2 相互独立，均服从密度函数为 $p(x)=\begin{cases}5e^{-5x}, & x>0\\ 0, & x\leqslant 0\end{cases}$ 的指数分布. 因此，可由独立随机变量和的密度公式求 T 的密度 $f(t)$，且可以根据独立随机变量和的性质求出 ET 和 DT.

[解] 以 X_1 和 X_2 表示先后开动的记录仪无故障工作的时间，则 $T=X_1+X_2$. 由条件知 $X_i\ (i=1,2)$ 的概率密度为

$$p_i(x)=\begin{cases}5e^{-5x}, & 若\ x>0,\\ 0, & 若\ x\leqslant 0.\end{cases}$$

两台仪器无故障工作时间 X_1 和 X_2 显然相互独立.

利用二独立随机变量和的密度公式求 T 的概率密度. 对于 $t>0$，有

$$f(t)=\int_{-\infty}^{+\infty}p_1(x)p_2(t-x)\mathrm{d}x = 25\int_0^t e^{-5x}e^{-5(t-x)}\mathrm{d}x$$
$$= 25e^{-5t}\int_0^t \mathrm{d}x = 25te^{-5t}.$$

当 $t\leqslant 0$ 时，显然 $f(t)=0$. 于是，得

$$f(t)=\begin{cases}25te^{-5t}, & 若\ t>0,\\ 0, & 若\ t\leqslant 0.\end{cases}$$

由于 X_i 服从参数为 $\lambda=5$ 的指数分布，知

$$EX_i=\frac{1}{5},\ DX_i=\frac{1}{25}\quad (i=1,2),$$

因此，有

$$ET=E(X_1+X_2)=EX_1+EX_2=\frac{2}{5}.$$

由于 X_1 和 X_2 独立，可见

$$DT=D(X_1+X_2)=DX_1+DX_2=\frac{2}{25}.$$

[典型错误] 不能由题意正确地概括出概率模型.

例 3.4.22 假设随机变量 Y 服从参数为 $\lambda=1$ 的指数分布. 随机变量

$$X_k=\begin{cases}0, & 若\ Y\leqslant k,\\ 1, & 若\ Y>k\end{cases}\quad (k=1,2).$$

(Ⅰ) 求 X_1 和 X_2 的联合概率分布；

(Ⅱ) 求 $E(X_1+X_2)$.

[提示] 本题主要考查 0－1 分布、指数分布、离散型随机变量和的分布和数学期望. X_1 和 X_2 分别以概率 $P\{Y>1\}$ 和 $P\{Y>2\}$ 服从 0－1 分布，是离散分布，$P\{Y>1\}$ 和 $P\{Y>2\}$ 由 Y 的分布给出. 求 (X_1,X_2) 的联合概率分布，需知道 (X_1,X_2) 的可能取值，显然只有四种可能. 在求每种可能取值的概率时要考虑它们与 Y 取值的关系，转化为连续型随机变量 Y 的概率计算问题. 根据随机变量和的性质知 $E(X_1+X_2)=EX_1+EX_2$，而 $EX_k=P\{Y>k\}$.

[解]（Ⅰ）Y 的分布函数为

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

(X_1,X_2)有四个可能值：(0,0)，(0,1)，(1,0)，(1,1).

易见

$$P\{X_1=0,X_2=0\}=P\{Y\leqslant 1,Y\leqslant 2\}=P\{Y\leqslant 1\}=1-e^{-1},$$

$$P\{X_1=0,X_2=1\}=P\{Y\leqslant 1,Y>2\}=0,$$

$$P\{X_1=1,X_2=0\}=P\{Y>1,Y\leqslant 2\}=P\{1<Y\leqslant 2\}=e^{-1}-e^{-2},$$

$$P\{X_1=1,X_2=1\}=P\{Y>1,Y>2\}=P\{X>2\}=e^{-2},$$

于是，可将 X_1 和 X_2 的联合概率分布列表如下：

X_2 \ P \ X_1	0	1
0	$1-e^{-1}$	$e^{-1}-e^{-2}$
1	0	e^{-2}

（Ⅱ）易见，$X_k\ (k=1,2)$服从 0－1 分布：

$$X_k\sim\begin{pmatrix}0 & 1\\ P\{Y\leqslant k\} & P\{Y>k\}\end{pmatrix}=\begin{pmatrix}0 & 1\\ 1-e^{-k} & e^{-k}\end{pmatrix},$$

因此，有

$$EX_k=P\{X_k=1\}=e^{-k}\ (k=1,2),$$

于是

$$E(X_1+X_2)=EX_1+EX_2=e^{-1}+e^{-2}.$$

[典型错误] 对 X_i 的定义不是很明白，故错误地认为 X_1 与 X_2 相互独立.

例 3.4.23 设两个随机变量 X，Y 相互独立，且都服从均值为 0、方差为$\frac{1}{2}$的正态分布，求随机变量$|X-Y|$的方差.

[提示] 本题主要考查正态分布、两个独立的随机变量函数的数字特征. 由于 X，Y 相互独立且服从正态分布，根据正态分布的重要性质，服从正态分布的随机变量的线性组合也服从正态分布. 令 $Z=X-Y$，则 Z 也服从正态分布，其均值 $EZ=EX-EY=0$．$DZ=DX+DY=1$．求 $D|X-Y|$关键是求$E|Z|$．因为

$$\begin{aligned}D|X-Y|&=D|Z|=E|Z|^2-(E|Z|)^2\\&=EZ^2-(E|Z|)^2=DZ+(EZ)^2-(E|Z|)^2\\&=1-(E|Z|)^2.\end{aligned}$$

由于 $Z\sim N(0,1)$，求 $E|Z|$实际上是计算一个简单的无穷限反常积分，这是比较容易的解法.

如果由 X、Y 的分布密度写出(X,Y)的联合分布密度，再计算二维随机变量函数$|X-Y|$的均值和方差，也能计算出来，但比较复杂.

[解法 1] 令 $Z=X-Y$．由于$X\sim N\left(0,\left(\frac{1}{\sqrt{2}}\right)^2\right)$，$Y\sim N\left(0,\left(\frac{1}{\sqrt{2}}\right)^2\right)$，且 X 和 Y 相互独立，故 $Z\sim N(0,1)$.

因为

$$D|X-Y|=D|Z|=E|Z|^2-(E|Z|)^2=EZ^2-(E|Z|)^2,$$

而

$$EZ^2=DZ=1,$$

$$E|Z|=\int_{-\infty}^{+\infty}|z|\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}ze^{-\frac{z^2}{2}}dz=\sqrt{\frac{2}{\pi}},$$

所以

$$D|X-Y|=1-\frac{2}{\pi}.$$

［解法 2］ X，Y 的概率密度分别为 $f_X(x)=\frac{1}{\sqrt{\pi}}e^{-x^2}$，$f_Y(y)=\frac{1}{\sqrt{\pi}}e^{-y^2}$．由于 X，Y 相互独立，故 X，Y 的联合概率密度函数为：

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

因此

$$\begin{aligned}E|X-Y|&=\frac{1}{\pi}\left[\iint\limits_{y<x}(x-y)e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y+\iint\limits_{y>x}(y-x)e^{-(x^2-y^2)}\mathrm{d}x\mathrm{d}y\right]\\&=\frac{1}{\pi}\left[\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}(x-y)e^{-(x^2+y^2)}\mathrm{d}y+\int_{-\infty}^{+\infty}\mathrm{d}y\int_{-\infty}^{y}(y-x)e^{-(x^2+y^2)}\mathrm{d}x\right]\\&=\frac{2}{\pi}\left[\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{x}(x-y)e^{-(x^2+y^2)}\mathrm{d}y\right]\\&=\frac{2}{\pi}\left(\int_{-\infty}^{+\infty}xe^{-x^2}\mathrm{d}x\int_{-\infty}^{x}e^{-y^2}\mathrm{d}y-\int_{-\infty}^{+\infty}e^{-x^2}\mathrm{d}x\int_{-\infty}^{x}ye^{-y^2}\mathrm{d}y\right)\\&=\frac{2}{\pi}\int_{-\infty}^{+\infty}e^{-2x^2}\mathrm{d}x\\&=\sqrt{\frac{2}{\pi}}.\end{aligned}$$

［典型错误］ 上述解法 2 利用到二重积分的技巧，计算过于繁琐，考生不仅花费大量宝贵时间，而且容易出错．有部分考生采用下面的典型错误解法：因为 X，Y 相互独立，故当 $X-Y>0$ 时，$D|X-Y|=D(X-Y)=DX+DY=1$；当 $X-Y<0$ 时，$D|X-Y|=D(Y-X)=DX+DY=1$．因此，$D|X-Y|=1$．这种解法是错误的．因为无论 $X-Y>0$，还是 $X-Y<0$．$D|X-Y|=DX+DY$ 都是不成立的．

例 3.4.24 设某种商品每周的需求量 X 是服从区间[10,30]上均匀分布的随机变量，而经销商店进货数量为区间[10,30]中的某一整数，商店每销售 1 单位商品可获利 500 元；若供大于求则削价处理，每处理 1 单位商品亏损 100 元；若供不应求，则可从外部调剂供应．此时每 1 单位商品仅获利 300 元．为使商店所获利润期望值不少于 9 280 元．试确定最少进货量．

［提示］ 本题主要考查均匀分布、随机变量的期望及一元二次函数的最值．由题意，X 的概率密度为 $f(x)=\begin{cases}\frac{1}{20}, & 20\leqslant x\leqslant 30,\\ 0, & 其他.\end{cases}$ 商店所获利润是需求量 X 和进货量 a 的函数，它也是随机变量的函数．题中已阐明供大于求和供不应求时利润与进货量的关系，关键是正确写出此函数关系，然后利用利润期望不少于 9 280 建立不等式解出 a 值．

［解］ 设进货数量为 a．则利润为

$$\begin{aligned}M_a&=\begin{cases}500a+(X-a)300, & a<X\leqslant 30,\\ 500X-(a-X)100, & 10\leqslant X\leqslant a\end{cases}\\&=\begin{cases}300X+200a, & a<X\leqslant 30,\\ 600X-100a, & 10\leqslant X\leqslant a.\end{cases}\end{aligned}$$

期望利润

$$\begin{aligned}EM_a&=\int_{10}^{30}\frac{1}{20}\cdot M_a\mathrm{d}x\\&=\frac{1}{20}\int_{10}^{a}(600x-100a)\mathrm{d}x+\frac{1}{20}\int_{a}^{30}(300x+200a)\mathrm{d}x\\&=\frac{1}{20}\left(600\cdot\frac{x^2}{2}-100ax\right)\Big|_{10}^{a}+\frac{1}{20}\left(300\cdot\frac{x^2}{2}+200ax\right)\Big|_{a}^{30}\\&=-7.5a^2+350a+5\ 250.\end{aligned}$$

依题意，有

$$-7.5a^2+350a+5\ 250\geqslant 9\ 280,$$

即
$$7.5a^2-350a+4\ 030\leqslant 0,$$

解得
$$\frac{62}{3}\leqslant a\leqslant 26.$$

故利润期望值不少于 9 280 元的最少进货量为 21 单位.

[典型错误] 写错 M_a 的表达式；计算 EM_a 时遗漏 X 的密度函数；解错不等式. 这些错误反映出部分考生的基本功欠扎实.

例 3.4.25 某箱装有 100 件产品，其中一、二和三等品分别为 80、10 和 10 件. 现在从中随机抽取一件. 记

$$X_i=\begin{cases}1, & \text{若抽到 } i \text{ 等品}, \\ 0, & \text{其他}\end{cases}\quad (i=1,2,3).$$

试求：(Ⅰ) 随机变量 X_1 与 X_2 的联合分布；

(Ⅱ) 随机变量 X_1 与 X_2 的相关系数 ρ.

[提示] 本题主要考查古典概率、随机变量的联合分布及相关系数. 由于 $X_i\ (i=1,2,3)$ 均只取 0 和 1 值，因此，二维随机变量 (X_1,X_2) 取值应为 $(0,0)$，$(0,1)$，$(1,0)$ 和 $(1,1)$. 需逐个计算它们的概率. 事件 $\{X_1=1,X_2=0\}$ 表示抽取的是一等品，$\{X_1=1,X_2=1\}$ 表示既是一等品又是二等品. 因此是不可能事件，等等. 由于相关系数 $\rho=\dfrac{\operatorname{cov}(X_1,X_2)}{\sqrt{DX_1\cdot DX_2}}$，因此，需分别求出 EX_1，EX_2，DX_1，DX_2 和 $E(X_1X_2)$.

[解] (Ⅰ) 设事件 $A_i=$ “抽到 i 等品” $(i=1,2,3)$.

由题意知 A_1，A_2，A_3 两两互不相容，且

$$P(A_1)=0.8,\ P(A_2)=P(A_3)=0.1,$$

易见，
$$P\{X_1=0,X_2=0\}=P(A_3)=0.1,$$
$$P\{X_1=0,X_2=1\}=P(A_2)=0.1,$$
$$P\{X_1=1,X_2=0\}=P(A_1)=0.8,$$
$$P\{X_1=1,X_2=1\}=P(\varnothing)=0.$$

(Ⅱ)
$$EX_1=0.8,\ EX_2=0.1,$$
$$DX_1=0.8\times 0.2=0.16,\ DX_2=0.1\times 0.9=0.09,$$
$$E(X_1X_2)=0\times 0\times 0.1+0\times 1\times 0.1+1\times 0\times 0.8+1\times 1\times 0=0,$$
$$\operatorname{cov}(X_1,X_2)=E(X_1X_2)-EX_1\cdot EX_2=0-0.8\times 0.1=-0.08,$$
$$\rho=\frac{\operatorname{cov}(X_1,X_2)}{\sqrt{DX_1\cdot DX_2}}=\frac{-0.08}{\sqrt{0.16\times 0.09}}=-\frac{2}{3}.$$

[典型错误] 将 X_1 和 X_2 看成是独立的，于是 $P\{X_1=1,X_2=1\}=P\{X_1=1\}\cdot P\{X_2=1\}=0.8\times 0.1=0.08$. 将 X_1，X_2 的联合分布写成：

X_1 \ P \ X_2	0	1
0	0.18	0.72
1	0.02	0.08

例 3.4.26 一商店经销某种商品，每周进货的数量 X 与顾客对该种商品的需求量 Y 是相互独立的随机变量，且都服从区间 $[10,20]$ 上的均匀分布. 商店每售出一单位商品可得利润 1 000 元；若需求量超过了进货量，商店可从其他商店调剂供应，这时每单位商品获利润为 500 元. 试计算此商店经销该种商品每周所得利润的期望值(如图 3.4.1 所示).

[提示] 本题主要考查均匀分布、随机变量的联合分布、随机变量的数学期望. 这是一道概率知识在经

济领域里的典型应用题. 前两句给出 X, Y 相互独立且都服从[10,20]上的均匀分布, 由此可写出(X,Y)的联合概率密度. 接下来两句阐明了可得利润与需求量和进货量之间的关系, 随机变量之间的函数关系是一个二元分段函数. 应用求二元随机变量函数均值定义求之, 将用到二重积分.

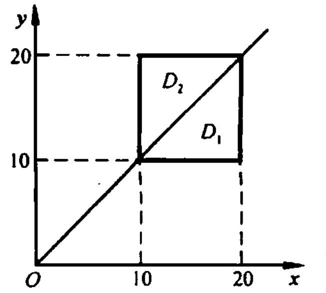

图 3.4.1

[解] 设 Z 表示商店每周所得的利润, 则

$$Z=\begin{cases}1\,000Y, & Y\leqslant X,\\ 1\,000X+500(Y-X)=500(X+Y), & Y>X.\end{cases}$$

由于 X 与 Y 的联合概率密度为

$$\varphi(x,y)=\begin{cases}\dfrac{1}{100}, & 10\leqslant x\leqslant 20,\ 10\leqslant y\leqslant 20,\\ 0, & 其他,\end{cases}$$

所以
$$\begin{aligned}EZ&=\iint\limits_{D_1}1\,000y\cdot\frac{1}{100}\mathrm{d}x\mathrm{d}y+\iint\limits_{D_2}500(x+y)\cdot\frac{1}{100}\mathrm{d}x\mathrm{d}y\\&=10\int_{10}^{20}\mathrm{d}y\int_{y}^{20}y\mathrm{d}x+5\int_{10}^{20}\mathrm{d}y\int_{10}^{y}(x+y)\mathrm{d}x\\&=10\int_{10}^{20}y(20-y)\mathrm{d}y+5\int_{10}^{20}\left(\frac{3}{2}y^2-10y-50\right)\mathrm{d}y\\&=\frac{20\,000}{3}+5\times1\,500\approx14\,166.67(元).\end{aligned}$$

[典型错误] 在计算过程中, 错误地将 Z 写成

$$Z=\begin{cases}1\,000Y, & Y\leqslant X,\\ 500Y, & Y>X.\end{cases}$$

在讨论需求量超过进货量$(Y>X)$的利润时, 忘却了已售完 X 的获利 $1\,000X$; 计算 EZ 时, 常将$\dfrac{1}{100}$遗漏; 计算 EZ 的后半部分纯属是计算二重积分, 有定错上、下限的, 有运算错的.

例 3.4.27 假设二维随机变量(X,Y)在矩形

$$G=\{(x,y)\mid 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$$

上服从均匀分布(如图 3.4.2 所示). 记为

$$U=\begin{cases}0, & 若\ X\leqslant Y,\\ 1, & 若\ X>Y.\end{cases}$$

$$V=\begin{cases}0, & 若\ X\leqslant 2Y,\\ 1, & 若\ X>2Y.\end{cases}$$

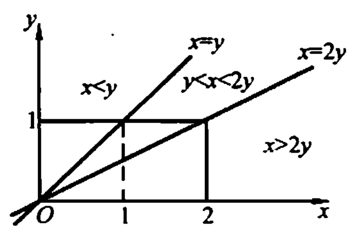

图 3.4.2

(Ⅰ) 求 U 和 V 的联合分布;

(Ⅱ) 求 U 和 V 的相关系数 ρ.

[提示] 本题主要考查二维均匀分布、离散型随机变量的联合分布、相关系数.

根据(X,Y)在矩形 G 上服从均匀分布, 很容易计算出 $P\{X\leqslant Y\}$, $P\{X\leqslant 2Y\}$的概率, 从而写出 U 和 V 的概率分布. U、V 均取 0 和 1 二值, 因此(U,V)只有(0,0), (0,1), (1,0)和(1,1)四种可能取值, 可根据 U, V 的分布求出它们, 从而也求出了 UV 的分布. 根据这些分布, 可较容易地求出有关变量的均值和方差以及 U 和 V 的协方差, 从而求出相关系数 r.

[解] 由题设可得

$$P\{X\leqslant Y\}=\frac{1}{4},$$

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

$$P\{Y<X\leqslant 2Y\}=\frac{1}{4}.$$

(Ⅰ) (U,V)有四个可能值：$(0,0)$，$(0,1)$，$(1,0)$，$(1,1)$.

$$P\{U=0,V=0\}=P\{X\leqslant Y,X\leqslant 2Y\}=P\{X\leqslant Y\}=\frac{1}{4}.$$

$$P\{U=0,V=1\}=P\{X\leqslant Y,X>2Y\}=0.$$

$$P\{U=1,V=0\}=P\{X>Y,X\leqslant 2Y\}=P\{Y<X\leqslant 2Y\}=\frac{1}{4}.$$

$$P\{U=1,V=1\}=1-\left(\frac{1}{4}+\frac{1}{4}\right)=\frac{1}{2}.$$

(Ⅱ) 由以上可见 UV 以及 U 和 V 的分布为

$$UV\sim\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix},\quad U\sim\begin{pmatrix}0 & 1\\ \frac{1}{4} & \frac{3}{4}\end{pmatrix},\quad V\sim\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}.$$

于是，有

$$EU=\frac{3}{4},\ DU=\frac{3}{16},\ EV=\frac{1}{2},\ DV=\frac{1}{4},\ E(UV)=\frac{1}{2},$$

$$\operatorname{cov}(U,V)=E(UV)-EU\cdot EV=\frac{1}{8},$$

$$\rho=\frac{\operatorname{cov}(U,V)}{\sqrt{DU\cdot DV}}=\frac{1}{\sqrt{3}}.$$

[典型错误] 不知如何计算 U，V 的联合概率分布及二者乘积的分布.

例 3.4.28 某流水生产线上每个产品不合格的概率为 $p\,(0<p<1)$，各产品合格与否相互独立．当出现一个不合格产品时即停机检修．设开机后第一次停机时已生产了的产品个数为 X．求 X 的数学期望 EX 和方差 DX．

[提示] 本题主要考查几何分布及其期望与方差．根据题意，开机后第一次停机时，若生产了的产品个数 $X=i$，则前 $i-1$ 个产品为合格产品，第 i 个产品必为不合格产品．因此，$P\{X=i\}=q^{i-1}p\ (q=1-p,i=1,2,\cdots)$，由此并根据数学期望和方差的定义求 EX 和 DX．

[解] 记 $q=1-p$．X 的概率分布为

$$P\{X=i\}=q^{i-1}p,i=1,2,\cdots.$$

X 的数学期望为

$$EX=\sum_{i=1}^{\infty}iq^{i-1}p=p\sum_{i=1}^{\infty}(q^i)'$$

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

因为

$$EX^2=\sum_{i=1}^{\infty}i^2q^{i-1}p=p\left[q\left(\sum_{i=1}^{\infty}q^i\right)'\right]'$$

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

所以 X 的方差为

$$DX=EX^2-(EX)^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

[典型错误]

① $P\{X=i\}=q^{i-1}p\ (i=1,2,\cdots)$写成 $P\{X=i\}=q^ip\ (i=1,2,\cdots)$，这是将合格产品的计数搞错了．如第一个产品就不合格，即该生产线未生产出合格产品就停机，因此应该是 $P\{X=1\}=q^0p$，从而知 $P\{X=i\}=q^{i-1}p\ (i=1,2,\cdots)$.

② 级数求和不会或求错，甚至有的写成有限项求和.

③ 有不少考生记住了这种分布(因为这是典型分布)的数学期望及方差的值，没有通过计算就将答案写上.

例 3.4.29 设随机变量 X 和 Y 的联合分布在以点(0,1)，(1,0)，(1,1)为顶点的三角形区域上服从均匀分布，试求随机变量 $U=X+Y$ 的方差.

[提示] 本题主要考查均匀分布、随机变量和的分布及其数字特征. 根据题意很容易写出(X,Y)的联合概率密度，可分别求出 X 和 Y 的边缘概率密度. 从而可以分别求出 EX，DX 和EY，DY（由对称性知，$EX=EY$，$DX=DY$，因此，求出其中一个即可）. 同时利用(X,Y)的联合概率密度可以求出$E(XY)$，从而求出 X，Y 的协方差 $\operatorname{cov}(X,Y)$，再求出方差 $DU=DX+DY+2\operatorname{cov}(X,Y)$. 当然也可以直接求出 $U=X+Y$ 的概率密度，然后求 DU.

[解法 1] 三角形区域为 $G=\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,x+y\geqslant 1\}$. 随机变量 X 和 Y 的联合密度为

$$f(x,y)=\begin{cases}2, & 若(x,y)\in G,\\ 0, & 若(x,y)\notin G.\end{cases}$$

以 $f_1(x)$表示 X 的概率密度，则当 $x\leqslant 0$ 或 $x\geqslant 1$ 时，$f_1(x)=0$；当 $0<x<1$ 时，有

$$f_1(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y=\int_{1-x}^{1}2\mathrm{d}y=2x,$$

因此

$$EX=\int_0^1 2x^2\mathrm{d}x=\frac{2}{3},\ EX^2=\int_0^1 2x^3\mathrm{d}x=\frac{1}{2},$$

$$DX=EX^2-(EX)^2=\frac{1}{2}-\frac{4}{9}=\frac{1}{18}.$$

同理可得 $EY=\frac{2}{3}$，$DY=\frac{1}{18}$.

现在求 X 和 Y 的协方差：

$$E(XY)=\iint_G 2xy\mathrm{d}x\mathrm{d}y=2\int_0^1 x\mathrm{d}x\int_{1-x}^{1}y\mathrm{d}y=\frac{5}{12},$$

$$\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY=\frac{5}{12}-\frac{4}{9}=-\frac{1}{36},$$

于是

$$\begin{aligned}DU&=D(X+Y)\\&=DX+DY+2\operatorname{cov}(X,Y)\\&=\frac{1}{18}+\frac{1}{18}-\frac{2}{36}=\frac{1}{18}.\end{aligned}$$

[解法 2] 三角形区域为 $G=\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,x+y\geqslant 1\}$，随机变量 X 和 Y 的联合密度为

$$f(x,y)=\begin{cases}2, & 若(x,y)\in G,\\ 0, & 若(x,y)\notin G.\end{cases}$$

以 $f(u)$表示 $U=X+Y$ 的概率密度.

当 $u<1$ 或 $u>2$ 时，显然 $f(u)=0$.

设 $1\leqslant u\leqslant 2$，当 $0\leqslant x\leqslant 1$ 且 $0\leqslant u-x\leqslant 1$ 时，$f(x,u-x)=2$，否则 $f(x,u-x)=0$. 由随机变量之和的概率密度公式，有

$$f(u)=\int_{-\infty}^{+\infty}f(x,u-x)\,\mathrm{d}x=\int_{u-1}^{1}2\mathrm{d}x=2(2-u),$$

因此

$$E(X+Y)=EU=\int_{-\infty}^{+\infty}uf(u)\,\mathrm{d}u=2\int_1^2 u(2-u)\,\mathrm{d}u=\frac{4}{3},$$

$$\begin{aligned}E(X+Y)^2=EU^2&=\int_{-\infty}^{+\infty}u^2f(u)\,\mathrm{d}u\\&=2\int_1^2 u^2(2-u)\,\mathrm{d}u=\frac{11}{6}.\end{aligned}$$

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

$$=\frac{11}{6}-\frac{16}{9}=\frac{1}{18}.$$

［典型错误］ 不会利用卷积公式求三角形区域上两随机变量和的概率密度函数.

例 3.4.30 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.

（Ⅰ）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；

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

［提示］ 本题主要考查事件独立概念、随机变量相关系数的性质. 题目给出了事件相关系数的定义，利用这一定义可直接证明（Ⅰ）. 由于任意两个随机变量 X，Y 的相关系数 ρ 满足 $|\rho|\leqslant 1$，故需考虑随机变量 X，Y 与事件 A，B 的概率的关系.

［证］ （Ⅰ）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.

（Ⅱ）考虑随机变量 X 和 Y：

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

由条件知，X 和 Y 都服从 0－1 分布：

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

易见

$$EX=P(A),\ EY=P(B),$$
$$DX=P(A)P(\bar{A}),\ DY=P(B)P(\bar{B}),$$
$$\operatorname{cov}(X,Y)=P(AB)-P(A)P(B),$$

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.

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

［典型错误］ 本题中，事件 A，B 的相关系数是在题目的条件中予以定义的. 多数考生能从这一定义出发顺利地证明（Ⅰ），但不少考生不知如何利用随机变量相关系数的性质去证明（Ⅱ）.

例 3.4.31 设 A，B 是二随机事件，随机变量

$$X=\begin{cases}1, & \text{若 } A \text{ 出现},\\ -1, & \text{若 } A \text{ 不出现};\end{cases}\qquad Y=\begin{cases}1, & \text{若 } B \text{ 出现},\\ -1, & \text{若 } B \text{ 不出现}.\end{cases}$$

试证明随机变量 X 和 Y 不相关的充分必要条件是 A 与 B 相互独立.

［提示］ 本题主要考查随机事件的独立性、随机变量的相关系数. 证明的目标是 $\operatorname{cov}(X,Y)=0\Leftrightarrow P(AB)=P(A)\cdot P(B)$，而 $\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY$，因而需正确写出 XY 的分布及 EX，EY 和 $E(XY)$.

［证］ 记 $P(A)=p_1$，$P(B)=p_2$，$P(AB)=p_{12}$. 由数学期望定义，可见

$$EX=P(A)-P(\bar{A})=2p_1-1,\ EY=2p_2-1.$$

现在求 $E(XY)$. 由于 XY 只有两个可能值 1 和 －1，可见

$$P\{XY=1\}=P(AB)+P(\bar{A}\bar{B})=2p_{12}-p_1-p_2+1,$$
$$P\{XY=-1\}=1-P\{XY=1\}=p_1+p_2-2p_{12},$$
$$E(XY)=P\{XY=1\}-P\{XY=-1\}=4p_{12}-2p_1-2p_2+1,$$

从而

$$\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY=4p_{12}-4p_1p_2,$$

因此,$\mathrm{cov}(X,Y)=0$ 当且仅当 $p_{12}=p_1p_2$，即 X 和 Y 不相关当且仅当事件 A 和 B 相互独立.

[典型错误] 不会用事件 A，B 的概率表示随机变量 X、Y 的联合分布.

例 3.4.32 设 $X_1,X_2,\cdots,X_n$ $(n>2)$为独立同分布的随机变量，且均服从 $N(0,1)$. 记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$Y_i=X_i-\overline{X}$，$i=1,2,\cdots,n$. 求：

(Ⅰ) Y_i 的方差 DY_i，$i=1,2,\cdots,n$；

(Ⅱ) Y_1 与 Y_n 的协方差 $\mathrm{cov}(Y_1,Y_n)$；

(Ⅲ) $P\{Y_1+Y_n\leqslant 0\}$.

[提示] 本题主要考查正态随机变量的期望与方差、相关性. 在求 DY_i 时应注意 X_i 与 $\overline{X}$ 不是独立的. 如果把 Y_i 改写成 $X_1,X_2,\cdots,X_n$ 的加权和，则可以用它们之间的独立性来计算 DY_i，$\mathrm{cov}(Y_1,Y_n)$及 Y_1+Y_n 的分布.

[解] (Ⅰ)
$$\begin{aligned}DY_i&=D(X_i-\overline{X})\\&=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum_{k\neq i}X_k\right]\\&=\frac{n-1}{n},\quad i=1,2,\cdots,n.\end{aligned}$$

(Ⅱ)
$$\begin{aligned}\mathrm{cov}(Y_1,Y_n)&=E(Y_1-EY_1)(Y_n-EY_n)\\&=E(X_1-\overline{X})(X_n-\overline{X})\\&=E(X_1X_n)+E\overline{X}^2-E(X_1\overline{X})-E(X_n\overline{X})\\&=EX_1EX_n+D\overline{X}-\frac{1}{n}EX_1^2-\frac{1}{n}\sum_{i=2}^{n}E(X_1X_i)\\&\quad-\frac{1}{n}EX_n^2-\frac{1}{n}\sum_{i=1}^{n-1}E(X_iX_n)\\&=-\frac{1}{n}.\end{aligned}$$

(Ⅲ)
$$\begin{aligned}Y_1+Y_n&=X_1-\overline{X}+X_n-\overline{X}\\&=\frac{n-2}{n}X_1-\frac{2}{n}\sum_{i=2}^{n-1}X_i+\frac{n-2}{n}X_n.\end{aligned}$$
上式是相互独立的正态随机变量的线性组合，所以 Y_1+Y_n 服从正态分布.

由于
$$E(Y_1+Y_n)=0,$$
故
$$P\{Y_1+Y_n\leqslant 0\}=\frac{1}{2}.$$

[典型错误] 误认为 X_i 与 $\overline{X}$ 是独立的，因而直接用 $DY_i=DX_i+D\overline{X}$ 来计算 DY_i.

例 3.4.33 设随机变量 X 的概率密度为
$$f_X(x)=\begin{cases}\frac{1}{2}, & -1<x<0,\\ \frac{1}{4}, & 0\leqslant x<2,\\ 0, & 其他,\end{cases}$$
令 $Y=X^2$，$F(x,y)$为二维随机变量(X,Y)的分布函数，求：

(Ⅰ) Y 的概率密度 $f_Y(y)$；

(Ⅱ) $\mathrm{cov}(X,Y)$；

(Ⅲ) $F\left(-\frac{1}{2},4\right)$.

[提示] 本题主要考查随机变量函数的概率密度的求法、数学期望和协方差的概念与性质及二维随机变量分布函数的定义.

[解] （Ⅰ）Y 的分布函数为

$$F_Y(y)=P\{Y\leqslant y\}=P\{X^2\leqslant y\}.$$

当 $y\leqslant 0$ 时，$F_Y(y)=0$，$f_Y(y)=0$；

当 $0<y<1$ 时，

$$\begin{aligned}F_Y(y)&=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}\\&=P\{-\sqrt{y}\leqslant X<0\}+P\{0\leqslant X\leqslant\sqrt{y}\}\\&=\frac{1}{2}\sqrt{y}+\frac{1}{4}\sqrt{y}=\frac{3}{4}\sqrt{y}.\end{aligned}$$

$$f_Y(y)=\frac{3}{8\sqrt{y}};$$

当 $1\leqslant y<4$ 时，

$$\begin{aligned}F_Y(y)&=P\{-1\leqslant X<0\}+P\{0\leqslant X\leqslant\sqrt{y}\}\\&=\frac{1}{2}+\frac{1}{4}\sqrt{y}.\end{aligned}$$

$$f_Y(y)=\frac{1}{8\sqrt{y}};$$

当 $y\geqslant 4$ 时，$F_Y(y)=1$，$f_Y(y)=0$.

故 Y 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{3}{8\sqrt{y}}, & 0<y<1,\\ \dfrac{1}{8\sqrt{y}}, & 1\leqslant y<4,\\ 0, & 其他.\end{cases}$$

（Ⅱ）

$$EX=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x=\int_{-1}^{0}\frac{1}{2}x\mathrm{d}x+\int_{0}^{2}\frac{1}{4}x\mathrm{d}x=\frac{1}{4}.$$

$$EY=EX^2=\int_{-\infty}^{+\infty}x^2f_X(x)\mathrm{d}x=\int_{-1}^{0}\frac{1}{2}x^2\mathrm{d}x+\int_{0}^{2}\frac{1}{4}x^2\mathrm{d}x=\frac{5}{6}.$$

$$E(XY)=EX^3=\int_{-\infty}^{+\infty}x^3f_X(x)\mathrm{d}x=\int_{-1}^{0}\frac{1}{2}x^3\mathrm{d}x+\int_{0}^{2}\frac{1}{4}x^3\mathrm{d}x=\frac{7}{8},$$

故

$$\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY=\frac{2}{3}.$$

（Ⅲ）
$$\begin{aligned}F\left(-\frac{1}{2},4\right)&=P\left\{X\leqslant-\frac{1}{2},Y\leqslant4\right\}=P\left\{X\leqslant-\frac{1}{2},X^2\leqslant4\right\}\\&=P\left\{X\leqslant-\frac{1}{2},-2\leqslant X\leqslant2\right\}=P\left\{-2\leqslant X\leqslant-\frac{1}{2}\right\}\\&=P\left\{-1<X\leqslant-\frac{1}{2}\right\}=\frac{1}{4}.\end{aligned}$$

[典型错误] 从试卷答题情况来看，有许多考生本题做得不理想，其主要原因是对概率密度为分段函数时，不知如何讨论，致使得出错误的结果．第(Ⅱ)问中大部分考生会用正确公式计算 EX、EY、$E(XY)$，$\operatorname{cov}(X,Y)$，但多数考生答案不对，这说明考生还需要加强计算能力．第(Ⅲ)问只需用二维随机变量分布函数的定义，再利用两事件的交进行计算，而相当一部分考生错误地利用随机变量的独立性概念做题．虽然答案正确，但做法不正确．还有的考生先求联合分布函数，复杂化了．该题较典型的错误有：

① 直接平方：

$$f_Y(y)=\begin{cases}\left(\dfrac{1}{2}\right)^2, & 0<y<1,\\ \left(\dfrac{1}{4}\right)^2, & 1\leqslant y<4,\\ 0, & 其他.\end{cases}$$

② 当 $y>0$ 时，

$$F_Y(y)=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}=\int_{-\sqrt{y}}^{\sqrt{y}}f_X(x)\mathrm{d}x$$

$$=\int_{-\sqrt{y}}^{0}\frac{1}{2}\mathrm{d}x+\int_{0}^{\sqrt{y}}\frac{1}{4}\mathrm{d}x=\frac{3}{4}\sqrt{y}.$$

没有分成 $0<y<1$，$1\leqslant y<4$ 来讨论.

③ 误认为

$$F\left(-\frac{1}{2},4\right)=P\left\{X\leqslant-\frac{1}{2},\ Y\leqslant4\right\}=P\left\{X\leqslant-\frac{1}{2}\right\}P\{Y\leqslant4\}$$

或

$$F\left(-\frac{1}{2},4\right)=\int_{-\infty}^{-\frac{1}{2}}\int_{-\infty}^{4}f(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{-\frac{1}{2}}f_X(x)\mathrm{d}x\cdot\int_{-\infty}^{4}f_Y(y)\mathrm{d}y.$$

五、大数定律和中心极限定理

大数定律是说明大量随机现象的平均值稳定性的一系列数学定理的总称，它有多种形式，常用的有伯努利大数定律和辛钦大数定律，切比雪夫大数定律是大数定律比较一般的形式.

中心极限定理揭示了“随机变量之和在一定条件下的极限分布是正态分布”，有着广泛的应用，也有着丰富的内容.

本节的一个重点是了解关于独立重复观测结果平均水平的极限分布是正态分布的列维－林德伯格中心极限定理和关于二项分布收敛于正态分布的棣莫弗－拉普拉斯中心极限定理；并会应用这两个定理近似计算有关事件的概率.

·考试内容与要求·

考试内容

切比雪夫大数定律　伯努利大数定律　辛钦(Khinchine)大数定律　棣莫弗－拉普拉斯(De Moivre－Laplace)定理　列维－林德伯格(Levy－Lindberg)定理

考试要求

1. 了解切比雪夫大数定律、伯努利大数定律和辛钦大数定律(独立同分布随机变量序列的大数定律).

2. 了解棣莫弗－拉普拉斯中心极限定理(二项分布以正态分布为极限分布)、列维－林德伯格中心极限定理(独立同分布随机变量序列的中心极限定理)，并会用相关定理近似计算有关随机事件的概率.

·考试内容解析·

(一) 依概率收敛

对于随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$，如果存在一个常数 A，对于任意给定的 $\varepsilon>0$，总有

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

则称随机变量列 $X_1,X_2,\cdots,X_n,\cdots$ 依概率收敛于 A，记作

$$X_n\xrightarrow{P}A\ 或\ P-\lim_{n\to\infty}X_n=A.$$

(二) 大数定律

1. 切比雪夫大数定律

设随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$ 两两独立或两两不相关，其期望与方差 EX_i，DX_i 均存在，且存在常数 $M>0$，使得 $DX_i\leqslant M$，$i=1,2,\cdots$，则对于任意的正数 ε，有

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

特别地，对于独立同分布的随机变量序列 $X_1,X_2,\cdots$(即对于任意的正整数 $k>1$，随机变量 $X_1,X_2,\cdots,X_k$ 相互独立且具有相同的概率分布)，若其期望 $EX_i=\mu$ 和方差 $DX_i=\sigma^2$ 存在，则对于任意的 $\varepsilon>0$，有