［典型错误］ 不能充分理解所给的已知条件及条件概率.

三、多维随机变量的分布

在许多问题中，一个随机试验的结果要用两个或两个以上的随机变量来描述，而每个随机变量的概率分布只能描述其自身的行为，它不能反映变量之间的联系与相互关系．多个随机变量的联合概率分布是描绘变量间统计相依关系的重要工具，通过它不仅可以描绘变量间的联合行为、性质和特征，还可以决定每个随机变量的概率分布.

多个随机变量的联合分布、边缘分布、相关性与独立性是概率论中的重要概念，也是数理统计的基础.本节的重点是：

(1) 理解多维随机变量的分布函数的基本概念和性质；离散型随机变量的联合分布、边缘分布和条件分布；连续型随机变量的联合概率密度、边缘密度和条件密度；会利用二维概率分布求有关事件的概率.

(2) 理解随机变量的独立性和不相关性的概念及联系，掌握随机变量独立的条件.

(3) 掌握二维均匀分布，了解二维正态分布的概率密度，理解其中参数的概率意义.

(4) 掌握根据两个随机变量的概率分布求其简单函数的概率分布的基本方法；会根据多个独立随机变量的概率分布求其函数的概率分布.

• 考试内容与要求 •

考试内容

多维随机变量及其分布函数　二维离散型随机变量的概率分布、边缘分布和条件分布　二维连续型随机变量的概率密度、边缘概率密度和条件密度　随机变量的独立性和不相关性　常见二维随机变量的分布　两个及两个以上随机变量的函数的分布

考试要求

1. 理解多维随机变量的分布函数的概念和基本性质.

2. 理解二维离散型随机变量的概率分布和二维连续型随机变量的概率密度，掌握二维随机变量的边缘分布和条件分布.

3. 理解随机变量的独立性和不相关性的概念，掌握随机变量相互独立的条件；理解随机变量的不相关性与独立性的关系.

4. 掌握二维均匀分布和二维正态分布，理解其中参数的概率意义.

5. 会根据两个随机变量的联合分布求其函数的分布，会根据多个相互独立随机变量的联合分布求其函数的分布.

• 考试内容解析 •

(一) 多维随机变量及其分布函数

1. 基本概念

(1) 联合分布函数：称 n 元函数

$$F(x_1,\cdots,x_n)=P\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\}$$

为 n 维随机变量(或随机向量)$X=(X_1,\cdots,X_n)$的分布函数或随机变量 $X_1,\cdots,X_n$ 的联合分布函数.

(2) 边缘分布函数：随机变量 $X_1,\cdots,X_n$ 中任意 k $(1\leqslant k<n)$个的联合分布函数，称为 $F(x_1,\cdots,x_n)$的边缘分布函数.

2. 联合分布的性质

以两个随机变量 X 和 Y 的联合分布函数 $F(x,y)$为例：

(1) $0\leqslant F(x,y)\leqslant 1$，且对每一个自变量单调不减.

(2) $F(x,y)$对于每一自变量右连续.

(3) $F(x,-\infty)=F(-\infty,y)=0$，$F(+\infty,+\infty)=1$.

(4) 对任意实数 $a<b$，$c<d$，有

$$P\{a<X\leqslant b,c<Y\leqslant d\}=F(b,d)-[F(b,c)+F(a,d)]+F(a,c).$$

(5) 随机变量 X 和 Y 的分布函数分别为

$$F_1(x)=F(x,+\infty),\quad F_2(y)=F(+\infty,y),$$

且它们完全由 X 和 Y 的联合分布函数 $F(x,y)$ 决定.

(二) 离散型随机变量的分布函数

每一个分量都是离散型随机变量的多维随机变量就称为多维离散型随机变量. 下面以二维为例讨论多维离散型随机变量的分布.

1. 联合概率分布

设 X 和 Y 均是离散型随机变量. 其可能的取值分别为 $\{x_i\}$ 和 $\{y_j\}$，则 X 和 Y 的联合概率分布表示为：

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,\ j=1,2,\cdots,$$

其中 $p_{ij}\geqslant 0$，$\sum\limits_i\sum\limits_j p_{ij}=1$.

2. 边缘概率分布

设 X 和 Y 的联合概率分布为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,\ j=1,2,\cdots,$$

则分别称

$$P\{X=x_i\}=\sum_j p_{ij}=p_{i\cdot},\quad i=1,2,\cdots,$$

$$P\{Y=y_j\}=\sum_i p_{ij}=p_{\cdot j},\quad j=1,2,\cdots$$

为 X 和 Y 关于其联合分布的边缘概率分布.

离散型随机变量的联合分布和边缘分布经常用下面的表格表示：

X \ p_{ij} \ Y	y_1	y_2	$\cdots$	y_n	$\cdots$	$p_{i\cdot}=\sum\limits_j p_{ij}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1n}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2n}	$\cdots$	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_m	p_{m1}	p_{m2}	$\cdots$	p_{mn}	$\cdots$	$p_{m\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$p_{\cdot j}=\sum\limits_i p_{ij}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot n}$	$\cdots$	1

3. 条件概率分布

设离散型随机变量 X 和 Y 的联合概率分布为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

对于给定的 j，如果 $P\{Y=y_j\}>0$，则称

$$P\{X=x_i\mid Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\cdots$$

为在条件 $Y=y_j$ 下的 X 的条件概率分布.

(三) 连续型随机变量的联合分布

每一个分量都是连续型随机变量的多维随机变量就称为多维连续型随机变量. 下面以二维为例讨论多维连续型随机变量的分布，而多维连续型随机变量的联合概率分布由概率密度表示.

1. 联合概率密度

设 X 和 Y 为二连续型随机变量，其联合分布函数为 $F(x,y)$，如果存在非负函数 $f(x,y)$，使得对任意实数 x 和 y，均有

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\mathrm{d}u\mathrm{d}v,$$

则称 $f(x,y)$ 为二维随机变量 (X,Y) 的概率密度或随机变量 X 和 Y 的联合概率密度.

2. 联合概率密度的性质

(1) $f(x,y)\geqslant 0$.

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

(3) 如果 $f(x,y)$ 在点 (x,y) 处连续，则

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}.$$

(4) (X,Y) 落在 xOy 平面上区域 G 内的概率为

$$P\{(X,Y)\in G\}=\iint_G f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

3. 边缘概率密度

随机变量 X 与 Y 的概率密度 $f_X(x)$ 和 $f_Y(y)$ 可以由其联合概率密度 $f(x,y)$ 得到：

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y,\quad f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x,$$

它们分别称为联合概率密度 $f(x,y)$ 的边缘概率密度.

4. 条件概率密度

设二维连续型随机变量 (X,Y) 的概率密度为 $f(x,y)$，X 的边缘概率密度为 $f_X(x)$. 若对于任意的 x，$f_X(x)>0$，则称

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)},\quad y\in\mathbf{R}$$

为 Y 在条件 $X=x$ 下的条件概率密度. 同样可以定义 X 在条件 $Y=y$ 下的条件概率密度 $f_{X|Y}(x|y)$.

另外，联合概率密度与条件概率密度和边缘概率密度之间有如下关系：

$$f(x,y)=f_X(x)\cdot f_{Y|X}(y|x)=f_Y(y)\cdot f_{X|Y}(x|y).$$

由条件概率密度得到下面的条件分布函数：

$$F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{f(t,y)}{f_Y(y)}\mathrm{d}t,$$

$$F_{Y|X}(y|x)=\int_{-\infty}^{y}\frac{f(x,t)}{f_X(x)}\mathrm{d}t.$$

(四) 随机变量的独立性

1. 一般概念

(1) 称随机变量 $X_1,\cdots,X_n$ 为相互独立的，如果其联合分布函数

$$F(x_1,\cdots,x_n)=F_1(x_1)\cdot\cdots\cdot F_n(x_n),$$

其中 $F_k(x_k)$，$k=1,\cdots,n$ 是随机变量 X_k 的分布函数.

(2) 称多维随机变量 $(X_1,\cdots,X_m)$ 与 $(X_{m+1},\cdots,X_n)$ 为相互独立的，如果

$$F(x_1,\cdots,x_m,x_{m+1},\cdots,x_n)=F_1(x_1,\cdots,x_m)\cdot F_2(x_{m+1},\cdots,x_n),$$

其中 $F_1(x_1,\cdots,x_m)$ 和 $F_2(x_{m+1},\cdots,x_n)$ 分别为 $(X_1,\cdots,X_m)$ 和 $(X_{m+1},\cdots,X_n)$ 的边缘分布函数.

(3) 称随机变量列 $X_1,\cdots,X_n,\cdots$ 是相互独立的，如果对于任意 $m\geqslant 2$，随机变量 $X_1,\cdots,X_m$ 相互独立.

2. 离散型随机变量的独立性

称离散型随机变量 $X_1,\cdots,X_n$ 相互独立，如果对于一切可能值 $x_1,\cdots,x_n$，有

$$P\{X_1=x_1,\cdots,X_n=x_n\}=P\{X_1=x_1\}\cdot\cdots\cdot P\{X_n=x_n\}.$$

3. 连续型随机变量的独立性

称连续型随机变量$X_1,\cdots,X_n$相互独立，如果它们的联合密度等于各变量密度之积：

$$f(x_1,\cdots,x_n)=f_{X_1}(x_1)\cdot\cdots\cdot f_{X_n}(x_n).$$

4. 性质

(1) 如果$X_1,\cdots,X_n$相互独立，则其中任意k $(2\leqslant k\leqslant n)$个随机变量也相互独立.

(2) 如果$X_1,\cdots,X_n$相互独立，则它们的函数$g_1(X_1)$，…，$g_n(X_n)$也相互独立.

(3) 如果两个随机变量独立，则一个变量关于另一个变量的条件分布就是其无条件分布.

(五) 常见的二维随机变量

1. 二维均匀分布

如果二维连续型随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{A}, & (x,y)\in G,\\ 0, & \text{其他},\end{cases}$$

则称X和Y的联合分布为区域G上的均匀分布，其中G为一平面有界区域，A为G的面积.

特别地，若$G=\{(x,y)\mid a\leqslant x\leqslant b,c\leqslant y\leqslant d\}$，则在区域$G$上的均匀分布密度为

$$f(x,y)=\begin{cases}\dfrac{1}{(b-a)(d-c)}, & \text{若 } a\leqslant x\leqslant b,\ c\leqslant y\leqslant d,\\ 0, & \text{其他}.\end{cases}$$

二维均匀分布的两个边缘分布、条件分布以及数字特征都与区域G的形状密切相关. 例如，$G=\{(x,y)\mid x^2+y^2\leqslant r^2\}$，则区域$G$上的二维均匀分布的两个边缘分布都不是均匀分布，而其中一个变量关于另一个变量的条件分布都是均匀分布. 再如$G=\{(x,y)\mid a\leqslant x\leqslant b,c\leqslant y\leqslant d\}$，则二维均匀分布的两个边缘分布分别为区间$(a,b)$和$(c,d)$内的均匀分布.

2. 二维正态分布

如果连续型随机变量X与Y的联合概率密度为

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

则称二维随机变量(X,Y)服从参数为μ_1，μ_2，σ_1，σ_2和ρ的二维正态分布，其中$\mu_1\in\mathbf{R},\mu_2\in\mathbf{R},\sigma_1>0,\sigma_2>0,\rho\in(-1,1)$为参数，且记为

$$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho).$$

对于二维正态分布，有如下性质：

(1) 二维正态分布的两个边缘分布均是正态分布：$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，但逆命题不成立：即便$X$和$Y$都服从正态分布，甚至$X$和$Y$的相关系数等于0，$X$和$Y$的联合分布也未必是二维正态分布.

(2) 若二维随机变量(X,Y)服从二维正态分布，则对于任给的实数a与b（至少有一个不为零），$aX+bY\sim N(a\mu_1+b\mu_2,a^2\sigma_1^2+b^2\sigma_2^2)$.

(3) 若二维随机变量(X,Y)服从二维正态分布，则相关系数$\rho=0$是X和Y独立的充分必要条件.

(六) 多个随机变量的函数的分布

包括多个独立随机变量之最大值、最小值的分布及其联合分布；关于两个随机变量之和、差、积、商的分布.

1. 最值的分布

设随机变量$X_1,\cdots,X_n$相互独立，分布函数分别为$F_1(x_1),\cdots,F_n(x_n)$，则随机变量$X_{(n)}=\max\{X_1,\cdots,X_n\}$与$X_{(1)}=\min\{X_1,\cdots,X_n\}$的分布函数分别为

$$F_{\max}(z)=\prod_{i=1}^{n}F_i(z),\quad F_{\min}(z)=1-\prod_{i=1}^{n}[1-F_i(z)].$$

特别地，若$X_1,\cdots,X_n$相互独立且有相同的分布函数$F(x)$，则有

$$F_{\max}(z)=[F(z)]^n,\quad F_{\min}(z)=1-[1-F(z)]^n.$$

2. 两个连续型随机变量和的分布

设二维连续型随机变量(X,Y)的概率密度为$f(x,y)$，则$Z=X+Y$的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\,\mathrm{d}x=\int_{-\infty}^{+\infty}f(z-y,y)\,\mathrm{d}y.$$

如果随机变量X与Y独立，则上式可化为

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\,\mathrm{d}x=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\,\mathrm{d}y.$$

其中$f_X(x)$和$f_Y(y)$分别为X与Y的边缘概率密度．上面的公式称为卷积公式．并记为f_X*f_Y，即

$$f_X*f_Y=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\,\mathrm{d}x=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\,\mathrm{d}y.$$

3. 两个随机变量函数的分布

设二维随机变量(X,Y)的概率密度为$f(x,y)$．则随机变量的函数$Z=g(X,Y)$的分布函数为

$$F_Z(z)=P\{Z\leqslant z\}=\iint\limits_{\{g(x,y)\leqslant z\}}f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

•例题详解•

例 3.3.1 设二维随机变量(X,Y)的概率密度为$f(x,y)=\begin{cases}6x, & 0\leqslant x\leqslant y\leqslant 1,\\ 0, & \text{其他},\end{cases}$则$P=\{X+Y\leqslant 1\}=$______．

[答案] $\frac{1}{4}$.

[提示] 本题主要考查二维随机变量由概率密度求概率的基本方法．方法选对，则问题转变成一个二重积分的计算．

[解]

$$\begin{aligned}P(X+Y\leqslant 1)&=\iint\limits_{x+y\leqslant 1}f(x,y)\,\mathrm{d}x\mathrm{d}y\\&=\int_0^{\frac{1}{2}}\mathrm{d}x\int_x^{1-x}6x\,\mathrm{d}y=\int_0^{\frac{1}{2}}6x(1-2x)\,\mathrm{d}x\\&=\frac{1}{4}.\end{aligned}$$

[典型错误] 将答案写成1，估计是这样得来的：

$$P\{X+Y\leqslant 1\}=\iint\limits_{x+y\leqslant 1}6x\,\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{1-x}6x\,\mathrm{d}y=1.$$

这是概念性的错误，$f(x,y)$是个分段函数，仅当$0\leqslant x\leqslant y\leqslant 1$时$f(x,y)=6x$，所以实际的积分区域应是$0\leqslant x\leqslant y\leqslant 1$与$x+y\leqslant 1$的交，而将所求概率直接写成$\iint\limits_{x+y\leqslant 1}6x\,\mathrm{d}x\mathrm{d}y$显然是错误的．

例 3.3.2 设平面区域D由曲线$y=\frac{1}{x}$及直线$y=0$，$x=1$，$x=e^2$所围成，二维随机变量(X,Y)在区域D上服从均匀分布，则(X,Y)关于X的边缘概率密度在$x=2$处的值为______．

[答案] $\frac{1}{4}$.

[提示] 本题主要考查二维均匀分布、二重积分及边缘概率密度．先由均匀分布的定义求出(X,Y)的联合概率密度，之后再求边缘概率密度即可．

[解] 设平面区域D的面积为S_D．由于二维随机变量(X,Y)在区域D上服从均匀分布．因此，(X,Y)的联合密度函数$f(x,y)$为

$$f(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D,\\ 0, & (x,y)\notin D.\end{cases}$$

而
$$S_D=\int_1^{e^2}dx\int_0^{\frac{1}{x}}dy=\int_1^{e^2}\frac{1}{x}dx=2,$$

$$f_X(x)=\int_0^{\frac{1}{x}}\frac{1}{2}dy=\frac{1}{2x},\ 1\leqslant x\leqslant e^2.$$

故 $f_X(2)=\frac{1}{4}$.

[典型错误] 不会求在 D 上均匀分布的概率密度函数.

例 3.3.3 设二维随机变量 (X,Y) 的概率分布为

$X \backslash Y$	0	1
0	0.4	a
1	b	0.1

若随机事件 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，则 $a=$ ______，$b=$ ______.

[答案] 0.4，0.1.

[提示] 本题主要考查事件的独立性、二维离散型随机变量的边缘分布及联合分布．首先由 $\{X=0\}$ 与 $\{X+Y=1\}$ 的独立性得一关于 a，b 的方程，之后再由边缘分布性质求得另一关于 a、b 的方程．联合两个方程求得 a，b 的值.

[解] 由已知可以求得 X、Y 的边缘分布为

$X \backslash Y$	0	1	和
0	0.4	a	$0.4+a$
1	b	0.1	$0.1+b$
和	$0.4+b$	$0.1+a$	1

由此可知

$$a+b=0.5 \qquad (*)$$

及
$$P\{X=0\}=0.4+a,\ P\{X+Y=1\}=a+b=0.5.$$

又由于 $\{X=0\}$ 与 $\{X+Y=1\}$ 独立，即

$$P\{X=0,X+Y=1\}=P\{X=0\}\cdot P\{X+Y=1\}.$$

故有
$$P\{X=0,Y=1\}=P\{X=0\}\cdot P\{X+Y=1\},$$

即
$$a=(0.4+a)\times 0.5,$$

由此求得 $a=0.4$，再由 $(*)$ 式可得 $b=0.1$.

[典型错误] 不清楚事件独立性的定义.

例 3.3.4 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 一定独立　　(B) (X,Y) 服从二维正态分布

(C) X 与 Y 未必独立　　(D) $X+Y$ 服从一维正态分布

[答案] (C).

[提示] 本题主要考查随机变量间的独立性与不相关性，特别是关于正态分布随机变量的独立性与相关性．对于正态分布，有下面的结论：①若 X 和 Y 的联合分布是二维正态分布，则 X 和 Y 独立的充分必要条

件是它们不相关. ②若 X 和 Y 都服从正态分布，且 X，Y 独立，则 (X,Y) 服从二维正态分布，$X+Y$ 服从一维正态分布.

［解］ 相互独立的两个随机变量 X，Y 一定不相关，反之未必成立. 故选项(A)不正确. 选项(B)，(D)不成立. 因为当 X，Y 都服从正态分布，且它们不相关时，(X,Y) 的联合分布未必是正态分布. 故选项(B)未必成立；$X+Y$ 也未必服从一维正态分布，选项(D)不正确. 故本题应选(C).

［典型错误］ 认为选项(A)或(D)正确. 其原因是混淆了二维正态分布与边缘分布为正态之间的区别.

例 3.3.5 设两个随机变量 X 与 Y 相互独立且同分布：$P\{X=-1\}=P\{Y=-1\}=\frac{1}{2}$，$P\{X=1\}=P\{Y=1\}=\frac{1}{2}$，则下列各式中成立的是(　　).

(A) $P\{X=Y\}=\frac{1}{2}$　　(B) $P\{X=Y\}=1$　　(C) $P\{X+Y=0\}=\frac{1}{4}$　　(D) $P\{XY=1\}=\frac{1}{4}$

［答案］ (A).

［提示］ 本题主要考查随机变量的独立性及二维随机变量的联合分布. 本题中的 X，Y 均是只取两个值的离散型随机变量，故四个选项中的每一个均可以通过直接计算而验证其正确性.

［解］ 因为 $P\{X=Y\}=P\{X=1,Y=1\}+P\{X=-1,Y=-1\}$

$$=\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2},$$

故选项(A)正确. 事实上，

$$P\{X+Y=0\}=\frac{1}{2},\ P\{XY=1\}=\frac{1}{2}.$$

［典型错误］ 不清楚随机变量相等的概念，而错选(B).

例 3.3.6 假设随机变量 X 服从指数分布，则随机变量 $Y=\min\{X,2\}$ 的分布函数(　　).

(A) 是连续函数　　(B) 至少有两个间断点　　(C) 是阶梯函数　　(D) 恰好有一个间断点

［答案］ (D).

［提示］ 本题主要考查指数分布、事件的基本运算. 只要利用已知信息，求出 $Y=\min\{X,2\}$ 的分布函数即可.

［解］ 令 $F_Y(y)=P\{Y<y\}=P\{\min\{X,2\}<y\}$，考虑 $y<2$ 和 $y\geqslant 2$ 两种情况. 当 $y<2$ 时，由 $\min\{X,2\}>y$，知 $X>y$，因而 $1-F_Y(y)=1-F_X(y)$，故 $F_Y(y)=F_X(y)=1-e^{-\lambda y}<1$；当 $y>2$ 时，$\{\min\{X,2\}<y\}$ 为必然事件，因此，$F_Y(y)=1$. $F_Y(y)$ 如图 3.3.1 所示. 因而，$F_Y(y)$ 恰有一个间断点.

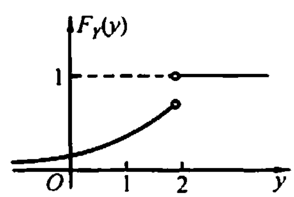

图 3.3.1

［典型错误］ 错误地选(A)，其原因可能是由于当 $y<2$ 时，$F_Y(y)=F_X(y)$ 所致.

例 3.3.7 设随机变量 X 在区间 $(0,1)$ 内服从均匀分布，在 $X=x\ (0<x<1)$ 的条件下，随机变量 Y 在区间 $(0,x)$ 内服从均匀分布. 求：

(Ⅰ) 随机变量 X 和 Y 的联合概率密度；

(Ⅱ) Y 的概率密度；

(Ⅲ) 概率 $P\{X+Y>1\}$.

［提示］ 本题主要考查条件分布、联合分布及边缘分布. 由已知得 $(Y|X=x)\sim U(0,x)$，$X\sim U(0,1)$，于是由条件概率密度公式可以求得 X，Y 的联合密度，之后再求边缘分布即可.

［解］ (Ⅰ) X 的概率密度为

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

在 $X=x\ (0<x<1)$ 的条件下，Y 的条件密度为

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

当 $0<y<x<1$ 时，随机变量 X 和 Y 的联合概率密度为

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=\frac{1}{x}.$$

在其他点 (x,y) 处，有 $f(x,y)=0$，即

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

（Ⅱ）当 $0<y<1$ 时，Y 的概率密度为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,dx=\int_y^1\frac{1}{x}dx=-\ln y;$$

当 $y\leqslant 0$ 或 $y\geqslant 1$ 时，$f_Y(y)=0$. 因此

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

（Ⅲ）所求概率

$$\begin{aligned}P\{X+Y>1\}&=\iint\limits_{x+y>1}f(x,y)\,dxdy\\&=\int_{\frac{1}{2}}^1dx\int_{1-x}^x\frac{1}{x}dy=\int_{\frac{1}{2}}^1\left(2-\frac{1}{x}\right)dx\\&=1-\ln 2.\end{aligned}$$

［典型错误］

① 不能由已知条件求出条件密度 $f_{Y|X}(y|x)$.

② 不知道关系式 $f(x,y)=f_X(x)f_{Y|X}(y|x)$.

③ 当 $x\leqslant 0$ 或 $x\geqslant 1$ 时，不会求 $f(x,y)$. 其实由 $f_X(x)=0$ 易得 $f(x,y)=0$.

④ 求出 $f_Y(y)=-\ln y$，而不是正确答案中的分段函数.

⑤ 在计算概率 $P\{X+Y>1\}$ 时，误为

$$P\{X+Y>1\}=\int_0^{\frac{1}{2}}dy\int_y^{1-y}\frac{1}{x}dx=\ln 2.$$

例 3.3.8 设某班车起点站上客人数 X 服从参数为 $\lambda\,(\lambda>0)$ 的泊松分布，每位乘客在中途下车的概率为 $p\,(0<p<1)$，且中途下车与否相互独立. 以 Y 表示在中途下车的人数. 求：

（Ⅰ）在发车时有 n 个乘客的条件下，中途有 m 人下车的概率；

（Ⅱ）二维随机变量 (X,Y) 的概率分布.

［提示］ 本题主要考查条件概率公式、独立性、二项分布、泊松分布、二维离散型随机变量的联合分布.

［解］ 问题（Ⅰ）是个条件概率，即当 $X=n$ 时，$Y=m$ 的概率 $P\{Y=m|X=n\}$. 由于车上的每位乘客下车与否是相互独立的，因此，Y 的条件概率分布为二项分布. 二维随机变量 (X,Y) 显然是离散的，根据条件概率公式知 $P\{X=n,Y=m\}=P\{Y=m|X=n\}\cdot P\{X=n\}$.

（Ⅰ）$P\{Y=m|X=n\}=C_n^mp^m(1-p)^{n-m},\quad 0\leqslant m\leqslant n,\ n=0,1,2,\cdots$.

（Ⅱ）$P\{X=n,Y=m\}=P\{Y=m|X=n\}P\{X=n\}$

$$=C_n^mp^m(1-p)^{n-m}\cdot\frac{e^{-\lambda}}{n!}\lambda^n,\quad 0\leqslant m\leqslant n,\ n=0,1,2,\cdots.$$

［典型错误］

① $(Y|X=n)\sim B(n,p)$ 写成 $Y\sim B(n,p)$.

② 不会计算二维离散型随机变量的分布.

③ 有部分考生不知“概率分布”是什么.

例 3.3.9　设随机变量 X 和 Y 的联合分布是正方形 $G=\{(x,y)\mid 1\leqslant x\leqslant 3,1\leqslant y\leqslant 3\}$ 上的均匀分布(如图 3.3.2 所示)，试求随机变量 $U=|X-Y|$ 的概率密度 $p(u)$.

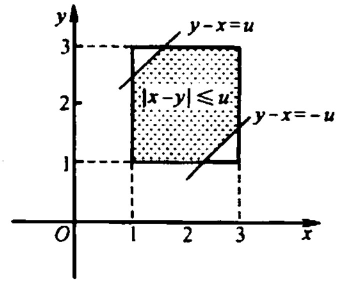

图 3.3.2

[提示]　本题主要考查均匀分布、随机变量函数的分布. 在求解时首先正确写出正方形 G 上的均匀分布的密度函数，然后利用分布函数的定义，先求出随机变量 U 的分布函数，求分布函数时注意将自变量 u 的变化区间进行合理的划分，以方便计算，最后求导便求出密度函数 $p(u)$.

[解]　由条件知 X 和 Y 的联合密度为

$$f(x,y)=\begin{cases}\frac{1}{4}, & 若 1\leqslant x\leqslant 3,\ 1\leqslant y\leqslant 3,\\ 0, & 其他.\end{cases}$$

以 $F(u)=P\{U\leqslant u\}$ $(-\infty<u<+\infty)$ 表示随机变量 U 的分布函数. 显然. 当 $u\leqslant 0$ 时，$F(u)=0$；当 $u\geqslant 2$ 时. $F(u)=1$；设 $0<u<2$，则

$$\begin{aligned}F(u)&=\iint\limits_{|x-y|\leqslant u} f(x,y)\mathrm{d}x\mathrm{d}y\\&=\iint\limits_{|x-y|\leqslant u}\frac{1}{4}\mathrm{d}x\mathrm{d}y\\&=\frac{1}{4}[4-(2-u)^2]=1-\frac{1}{4}(2-u)^2.\end{aligned}$$

于是，随机变量的概率密度为

$$p(u)=\begin{cases}\frac{1}{2}(2-u), & 若 0<u<2,\\ 0, & 其他.\end{cases}$$

[典型错误]　U 的取值范围及二重积分的计算.

例 3.3.10　设随机变量 X 与 Y 独立，其中 X 的概率分布为

$$X\sim\begin{pmatrix}1 & 2\\ 0.3 & 0.7\end{pmatrix},$$

而 Y 的概率密度为 $f(y)$，求随机变量 $U=X+Y$ 的概率密度 $g(u)$.

[提示]　本题主要考查随机变量的独立性、离散型随机变量、连续型随机变量、全概率公式. 要解答该题，实际需要解决两个问题：①要说明 $U=X+Y$ 是连续型随机变量. ②求 U 的密度函数. 解决这两个问题都需要知道随机变量 U 的分布函数，在求 U 的分布函数过程中要用全概率公式及随机变量独立的性质，然后再应用连续型随机变量的概率密度与其分布函数的关系导出 $g(u)$.

[解]　设 $F(y)$ 是 Y 的分布函数，则由全概率公式知 $U=X+Y$ 的分布函数为：

$$\begin{aligned}G(u)&=P\{X+Y\leqslant u\}\\&=0.3P\{X+Y\leqslant u\mid X=1\}+0.7P\{X+Y\leqslant u\mid X=2\}\\&=0.3P\{Y\leqslant u-1\mid X=1\}+0.7P\{Y\leqslant u-2\mid X=2\}.\end{aligned}$$

由于 X 与 Y 独立，可见

$$G(u)=0.3P\{Y\leqslant u-1\}+0.7P\{Y\leqslant u-2\},$$

由此得 U 的概率密度

$$\begin{aligned}g(u)=G'(u)&=0.3F'(u-1)+0.7F'(u-2)\\&=0.3f(u-1)+0.7f(u-2).\end{aligned}$$

[典型错误]　直接用卷积公式是一种错误，因为 X 与 Y 是两个不同类型的随机变量. 另外，也有部分考生不会正确利用全概率公式.

例 3.3.11　设随机变量 X 与 Y 相互独立，下表列出了二维随机变量 (X,Y) 的联合分布律及关于 X 和 Y 的边缘分布律中的部分数值，试将其余数值填入表中的空白处.

X \ Y	y_1	y_2	y_3	$P\{X=x_i\}=p_{i\cdot}$
x_1		$\frac{1}{8}$		
x_2	$\frac{1}{8}$			
$P\{Y=y_j\}=p_{\cdot j}$	$\frac{1}{6}$			1

［提示］ 本题主要考查二维离散型随机变量的联合分布、边缘分布的性质．由于 X、Y 相互独立，因此

$$P\{X=x_i, Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\},\quad i,j=1,2.$$

根据边缘概率分布的性质还知：

$$P\{Y=y_j\}=P\{X=x_1, Y=y_j\}+P\{X=x_2, Y=y_j\}$$

以及

$$P\{X=x_i\}=P\{X=x_i, Y=y_1\}+P\{X=x_i, Y=y_2\}+P\{X=x_i, Y=y_3\}.$$

若记 $p_{ij}=P\{X=x_i, Y=y_j\}$，则

$$p_{11}=\frac{1}{6}-\frac{1}{8}=\frac{1}{24},\quad p_{1\cdot}=\frac{p_{11}}{p_{\cdot 1}}=\frac{1}{4},\quad p_{13}=p_{1\cdot}-p_{11}-p_{12}=\frac{1}{4}-\frac{1}{24}-\frac{1}{8}=\frac{1}{12}.$$

同理，可计算出 $p_{2\cdot}$，$p_{\cdot 2}$，$p_{\cdot 3}$，p_{22}，p_{23}．

［解］

X \ Y	y_1	y_2	y_3	$P\{X=x_i\}=p_{i\cdot}$
x_1	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{4}$
x_2	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{4}$
$P\{Y=y_j\}=p_{\cdot j}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$	1

［典型错误］ 不知利用边缘分布与联合分布之间的关系而无法完成该题．

例 3.3.12 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 3.3.3 所示），试求边长为 X 和 Y 的矩形面积 S 的概率密度 $f(s)$．

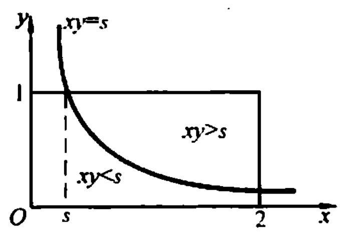

图 3.3.3

［提示］ 本题主要考查二维均匀分布、随机变量乘积的概率分布．首先作出矩形 G 的图形，并写出 (X, Y) 的密度函数 $\psi(x,y)$．$S=XY$ 为二维随机变量 (X, Y) 的函数，显然，只要求出 S 的分布函数 $F(s)=P\{XY\leqslant s\}$ 即可．

［解］ 二维随机变量 (X, Y) 的概率密度为

$$\psi(x,y)=\begin{cases}\frac{1}{2}, & 若(x,y)\in G,\\ 0, & 若(x,y)\overline{\in} G.\end{cases}$$

设 $F(s)=P\{S\leqslant s\}$ 为 S 的分布函数，则：

当 $s\leqslant 0$ 时，$F(s)=0$；

当 $s\geqslant 2$ 时，$F(s)=1$．

现在，设 $0<s<2$．曲线 $xy=s$ 与矩形 G 的上边交于点 $(s,1)$；位于曲线 $xy=s$ 上方的点满足 $xy>s$，位

于下方的点满足 $xy<s$，于是

$$F(s)=P\{S\leqslant s\}=P\{XY\leqslant s\}=1-P\{XY>s\}$$
$$=1-\iint\limits_{xy>s}\frac{1}{2}\mathrm{d}x\mathrm{d}y=1-\frac{1}{2}\int_{s}^{2}\mathrm{d}x\int_{\frac{s}{x}}^{1}\mathrm{d}y$$
$$=\frac{s}{2}(1+\ln 2-\ln s),$$

故

$$f(s)=\begin{cases}\frac{1}{2}(\ln 2-\ln s), & 若\ 0<s<2,\\ 0, & 若\ s\leqslant 0\ 或\ s\geqslant 2.\end{cases}$$

[典型错误] 不会利用联合分布求随机变量函数的概率分布．另外，在求概率 $P\{XY\leqslant s\}$时，不会划分积分区域．

例 3.3.13 已知随机变量 X_1 和 X_2 的概率分布

$$X_1\sim\begin{pmatrix}-1 & 0 & 1\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{pmatrix},\quad X_2\sim\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix},$$

而且 $P\{X_1X_2=0\}=1$.

（Ⅰ）求 X_1 和 X_2 的联合分布；

（Ⅱ）问 X_1 和 X_2 是否独立？为什么？

[提示] 本题主要考查二维随机变量的联合分布及独立性．由于 $P\{X_1X_2=0\}=1$，则 $P\{X_1X_2\neq 0\}=0$，因此，$P\{X_1=-1,X_2=1\}=P\{X_1=1,X_2=1\}=0$．这样就可根据边缘分布的性质，立即写出 X_1，X_2 的联合分布．至于第(Ⅱ)问，只要找到一个例子说明

$$P\{X_1=x_1,X_2=x_2\}\neq P\{X_1=x_1\}\cdot P\{X_2=x_2\},$$

则 X_1 和 X_2 不独立，否则独立．

[解法 1]（Ⅰ）由 $P\{X_1X_2=0\}=1$．可见

$$P\{X_1=-1,X_2=1\}=P\{X_1=1,X_2=1\}=0.$$

易见

$$P\{X_1=-1,X_2=0\}=P\{X_1=-1\}=\frac{1}{4},$$
$$P\{X_1=0,X_2=1\}=P\{X_2=1\}=\frac{1}{2},$$
$$P\{X_1=1,X_2=0\}=P\{X_1=1\}=\frac{1}{4},$$
$$P\{X_1=0,X_2=0\}=1-\left(\frac{1}{4}+\frac{1}{2}+\frac{1}{4}\right)=0,$$

于是，得 X_1 和 X_2 的联合分布

X_2 \ X_1	-1	0	1	Σ
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
Σ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

（Ⅱ）由以上结果，可见

$$P\{X_1=0,X_2=0\}=0.$$

$$P\{X_1=0\}P\{X_2=0\}=\frac{1}{4}\neq 0.$$

于是，X_1 和 X_2 不独立.

[解法 2] （Ⅰ）由 $P\{X_1X_2=0\}=1$，可见

$$P\{X_1=-1,X_2=1\}=P\{X_1=1,X_2=1\}=0,$$

因此，X_1 和 X_2 的联合分布有如下结构：

X_2 \ X_1	-1	0	1	Σ
0	p_{11}	p_{21}	p_{31}	$\frac{1}{2}$
1	0	p_{22}	0	$\frac{1}{2}$
Σ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

于是，由上表易见 X_1 和 X_2 的联合分布为：

X_2 \ X_1	-1	0	1	Σ
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
Σ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

（Ⅱ）同解法 1.

[典型错误] 不会正确利用已知条件 $P\{X_1X_2=0\}=1$.

例 3.3.14 设二维随机变量 (X,Y) 的概率密度为

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

求：（Ⅰ）(X,Y) 的边缘概率密度 $f_X(x)$，$f_Y(y)$；

（Ⅱ）$Z=2X-Y$ 的概率密度 $f_Z(z)$；

（Ⅲ）$P\left\{Y\leqslant\frac{1}{2}\,\middle|\,X\leqslant\frac{1}{2}\right\}$.

[提示] 本题主要考查二维随机变量的边缘概率密度、随机向量函数的概率密度函数、条件概率. 在由联合概率密度计算边缘概率密度时，应注意积分区域. 在求 $Z=2X-Y$ 的概率密度时可以用卷积公式，也可以直接用 (X,Y) 的联合概率密度.

[解法 1] （Ⅰ）当 $0<x<1$ 时，

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y=\int_0^{2x}\mathrm{d}y=2x;$$

当 $x\leqslant 0$ 或 $x\geqslant 1$ 时，$f_X(x)=0$，故

$$f_X(x)=\begin{cases}2x, & 0<x<1,\\ 0, & 其他.\end{cases}$$

当 $0<y<2$ 时，

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

当 $y\leqslant0$ 或 $y\geqslant2$ 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 $z\leqslant0$ 时，$F_Z(z)=0$；
当 $z\geqslant2$ 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) $P\left\{Y\leqslant\dfrac{1}{2}\,\middle|\,X\leqslant\dfrac{1}{2}\right\}=\dfrac{P\left\{X\leqslant\dfrac{1}{2},\ Y\leqslant\dfrac{1}{2}\right\}}{P\left\{X\leqslant\dfrac{1}{2}\right\}}=\dfrac{3/16}{1/4}=\dfrac{3}{4}.$

[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 $z\leqslant0$ 或 $z\geqslant2$ 时，$f_Z(z)=0$；

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

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