第 3 讲　二维随机变量的概率分布

知识梳理

一　二维离散型随机变量的分布律

1.联合分布律　p_{ij}

X，Y 的联合分布律

$$P(X=x_i,Y=y_j)=p_{ij}$$

性质：$p_{ij}\geq 0$　且　$\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty}p_{ij}=1$

2.边际分布律　p_i　p_j

X 的边际分布律

$$P(X=x_i)=\sum_{j=1}^{\infty}p_{ij}=p_i$$

Y 的边际分布律

$$P(Y=y_j)=\sum_{i=1}^{\infty}p_{ij}=p_j$$

3.条件分布律　p_{ij}/p_i　p_{ij}/p_j

X 的条件分布律

$$P(X=x_i|Y=y_j)=\frac{p_{ij}}{p_j}$$

Y 的条件分布律

$$P(Y=y_j|X=x_i)=\frac{p_{ij}}{p_i}$$

二　二维随机变量的分布函数

1.联合分布函数　$F(x,y)$

联合分布函数

$$F(x,y)=P(X\leq x,Y\leq y)$$

2.边际分布函数　$F_X(x)$　$F_Y(y)$

X 的边际分布函数

$$F_X(x)=F(x,+\infty)$$

Y 的边际分布函数

$$F_Y(y)=F(+\infty,y)$$

3.条件分布函数　$F_{X|Y}(x\,|\,y_j)$　$F_{Y|X}(y\,|\,x_i)$

X 的条件分布函数

$$F_{Y|X}(y\,|\,x_i)=P(Y\leq y\,|\,X=x_i)$$

Y 的条件分布函数

$$F_{X|Y}(x\,|\,y_j)=P(X\leq x\,|\,Y=y_j)$$

三　二维连续型随机变量的密度函数

1.联合密度函数　$f(x,y)$

联合密度函数

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

性质：$f(x,y)\geq 0$　且　$\iint\limits_{x,y\in(-\infty,+\infty)}f(x,y)\mathrm{d}x\mathrm{d}y=1$

· 通过在区域 D 内积分来获得 (X,Y) 落在某区域 D 内的概率

二维连续型概率求解

$$P((X,Y)\in D)=\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$$

2.**边际密度函数** $f_X(x)$ $f_Y(y)$

X 的边际密度函数

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

Y 的边际密度函数

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

3.**条件密度函数** $f_{X|Y}(x\mid y)$ $f_{Y|X}(y\mid x)$

X 的条件密度函数

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

Y 的条件密度函数

$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)}$$

四　二维随机变量的特征

1.**独立性**

二维随机变量独立性判断

二维随机变量 X,Y 相互独立 $\Leftrightarrow$ $F(x,y)=F_X(x)F_Y(y)$ $p_{ij}=p_ip_j$ 或 $f(x,y)=f_X(x)f_Y(y)$

2.**数学期望**

离散型数学期望

$$E(h(X,Y))=\sum_{i=1}^{+\infty}h(x_i,y_j)p_{ij}$$

连续型数学期望

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

3.**协方差**

协方差

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

4.**相关系数**

相关系数

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

5.**独立与不相关之间的关系**

- 独立和不相关看似是等价概念，但实际上这个相关指的是线性相关

定理　独立 $\Rightarrow$ 不相关　不相关 $\nRightarrow$ 独立

题型解析

六　求解二元离散型非常见随机变量分布

1.题型简述与解法

① 根据实际背景求联合分布律　例 1 (1)

· 一般使用二维表格来表示联合分布律，并额外加一行一列记录边际分布律

$i\,/\,j$	Y 的取值				p_i
X 的取值	p_{11}	p_{12}	$\cdots\cdots$	p_{1n}	X 边际分布律
	p_{21}	p_{22}	$\cdots\cdots$	p_{2n}	
	$\cdots\cdots$	$\cdots\cdots$	$\cdots\cdots$	$\cdots\cdots$	
	p_{m1}	p_{m2}	$\cdots\cdots$	p_{mn}	
p_j	Y 的边际分布律				$\cdots\cdots$

· 如果是根据实际背景，先列出 X 和 Y 的所有可能取值 $x_1,\cdots,x_m$， $y_1,\cdots,y_n$， 一一组合，列出表格然后依次求出 $P(X=x_i,Y=y_j)$，填入表格，得到分布律

· 不要漏掉可能的取值情况，计算量大，不要出错，算完后可以看看加起来是否等于 1

② 根据联合分布律求事件概率、期望、协方差、相关系数　例 1 (2)

· 求事件概率时，找出所有满足该事件的取值 (X,Y)，然后把对应的概率相加

· 求 $g(X,Y)$ 的数学期望时，先在联合分布律表格中计算并标记各个 (X,Y) 对应的 $g(X,Y)$ 值然后再列式 $E(g(X,Y))=\cdots$ 计算，如果 $g(X,Y)$ 为 0，可以直接跳过这一项，减少运算时间

③ 根据已知条件求联合分布律　例 2

· 自带的已知条件：同行同列相加为边际分布律、边际分布律之和为 1

· 题目可能提供的条件：期望、协方差、相互独立等

· 解法：用未知数表示联合分布律，利用知识点中的公式表示出已知信息，联立求解

· X 与 Y 相互独立 $\Leftrightarrow$ 所有的 $P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$

2.历年考试典型例题

例 1　（14—15 **秋冬　节选**）盒中有 3 个红球，5 个白球，从中随机取一球，观察其颜色后放回，并从别处拿两个与取出的球同色的球放入盒中，搅匀后再从中取一球，$X_i=\begin{cases}1, & 第\,i\,次取到红球\\ 0, & 第\,i\,次取到白球\end{cases}$，$i=1,2$。

求：（1）(X_1,X_2) 的联合分布律及 X_2 的边际分布律；（2）求 X_1 与 X_2 的相关系数 $\rho_{X_1X_2}$.

解　（1）两个随机变量为 X_1 和 X_2，根据题干描述的背景，可以得到

$$P(X_1=1)=\frac{3}{8},\quad P(X_1=0)=\frac{5}{8}$$

如果第一次摸到的是红球，则第二次共有 5 红球 5 白球，此时摸到红球的概率为 1/2，即

$$P(X_2=1|X_1=1)=\frac{1}{2}, \quad P(X_2=0|X_1=1)=\frac{1}{2}$$

如果第一次摸到的是白球，则第二次共有 3 红球 7 白球，此时摸到红球的概率为 3/10，即

$$P(X_2=1|X_1=0)=\frac{3}{10}, \quad P(X_2=0|X_1=0)=\frac{7}{10}$$

因此，通过条件概率求出联合分布律（如 $P(X_1=1,X_2=1)=P(X_1=1)P(X_2=1|X_1=1)$）

将每一列的概率相加就得到 X_2 的边际分布律：

$X_1 \backslash X_2$	0	1	$P(X_1=i)$
0	7/16	3/16	5/8
1	3/16	3/16	3/8
$P(X_2=j)$	5/8	3/8	

（2）根据求得的分布律为：

$X_1 \backslash X_2$	0	1	$P(X_1=i)$
0	7/16 0	3/16 0	5/8 0
1	3/16 0	3/16 1	3/8 1
$P(X_2=j)$	5/8 0	3/8 1	

→ $E(X_1X_2)=1\times\frac{3}{16}=\frac{3}{16}$（见标红） $E(X_1)=E(X_2)=1\times\frac{3}{8}=\frac{3}{8}$

→ $E(X_1^2)=E(X_2^2)=1\times\frac{3}{8}=\frac{3}{8}$（见标蓝） → $D(X_1)=D(X_2)=\frac{3}{8}-\left(\frac{3}{8}\right)^2=\frac{15}{64}$

→ $\mathrm{Cov}(X_1,X_2)=\frac{3}{16}-\left(\frac{3}{8}\right)^2=\frac{3}{64}$ → $\rho_{X_1X_2}=\frac{\mathrm{Cov}(X_1,X_2)}{\sqrt{D(X_1)D(X_2)}}=\frac{3/64}{15/64}=\boxed{1/5}$

例 2 （**17—18 春夏**）设 (X,Y) 的联合分布律如下表所示.已知 $E(X)=0.6$， $E(Y)=0$.

$X \backslash Y$	-1	0	1
0	0.1	a_2	a_3
1	a_4	a_5	a_6

（1）若 $a_6=0.1$，且 X 与 Y 不相关，求 (X,Y) 的联合分布律；

（2）若 X 与 Y 相互独立，求 (X,Y) 的联合分布律.

解 （1）由 X 与 Y 不相关 → $E(XY)=E(X)E(Y)=0$

$E(XY)=-1\times a_4+1\times a_6=a_6-a_4=0$（见下表标红） → $a_4=a_6=0.1$

$X \backslash Y$	-1	0	1
0	0.1 0	a_2 0	a_3 0
1	a_4 -1	a_5 0	a_6 1

∴ $E(X)=a_4+a_5+a_6=0.6$ → $a_5=0.4$

$E(Y)=-(0.1+a_4)+a_3+a_6=0$ → $a_3=0.1$

$a_2+\cdots+a_6=0.9 \quad \to \quad a_2=0.1$，联合分布律如下：

$X\backslash Y$	-1	0	1
0	0.1	0.1	0.1
1	0.1	0.4	0.1

（2）由 $E(X)=1\cdot P(X=1)=0.6$，得 $P(X=1)=0.6$，$P(X=0)=0.4$：

$X\backslash Y$	-1	0	1	
0	0.1	a_2	a_3	0.4
1	a_4	a_5	a_6	0.6

由于相互独立，此时 $P(X=0,Y=-1)=P(X=0)P(Y=-1)=0.4P(Y=-1)=0.1$

$\therefore \quad P(Y=-1)=0.25$

$\because \quad E(Y)=P(Y=1)-P(Y=-1)=0 \ \to \ P(Y=1)=0.25$，$P(Y=0)=0.5$

由此边际分布律全部求得，可求出联合分布律如下表所示

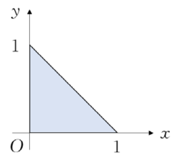
$X\backslash Y$	-1	0	1	
0	0.1	0.2	0.1	0.4
1	0.15	0.3	0.15	0.6
	0.25	0.5	0.25	

七　求解二元连续型非常见随机变量分布

1. 题型简述与解法

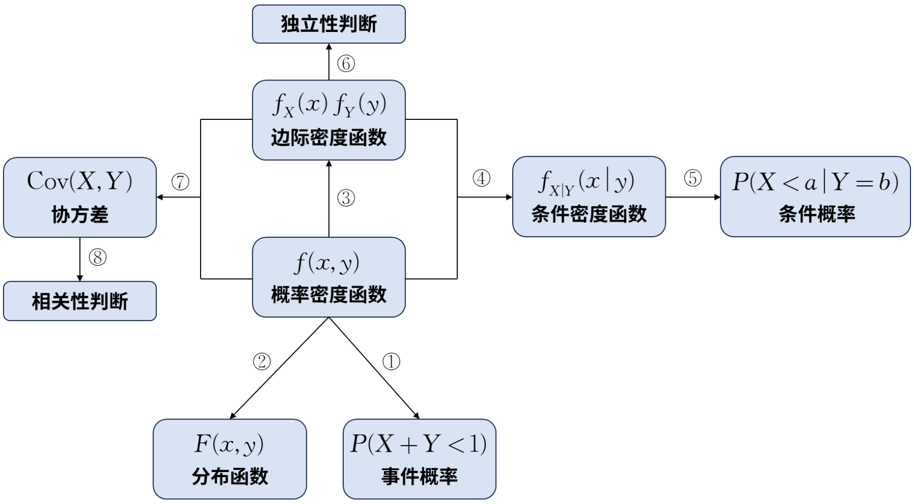
- 从二元随机变量 (X,Y) 的联合密度函数出发，计算、判断多种特征
- 此类题型首先应**画出概率密度函数的非零区域**，如下面这个函数的非零区域就是一个等腰直角三角形

$$f(x,y)=\begin{cases}6x, x>0, y>0, x+y<1\\0, \quad \text{其他}\end{cases} \quad \to$$



有了这张图，我们在后续积分时思路会更加清晰，不易出错

- 计算的总体思路如下图所示：



① **求事件概率** 例 1 (1)

· 确认该事件对应的区域，与我们画好的非零区域取交集，得到要积分的区域，在该区域内积分即可

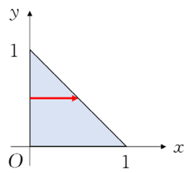
② **求分布函数的值** 例 2 (1)

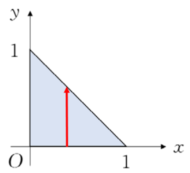
· 将分布函数转化为对应的事件，按①处理

③ **求边际密度函数** 例 1 (2) 例 2 (2)

· 代入边际密度函数的定义式进行积分

若 $f(x,y)$ 只在有限区域 D 内不为 0，则只在不为 0 的 x 或 y 上积分即可





求 $f_Y(y)$：对 x 从左往右积分　　求 $f_X(x)$：对 y 从下往上积分

· 最终的结果要求 $x,y\in(-\infty,+\infty)$，**因此要写成分段函数的形式**（加上 0，其他）

④ **求条件密度函数** 例 1 (3) 例 2 (3)

· 求出边际密度后，代入条件密度函数的定义式即可

· 此时也要注意 x 或 $y\in(-\infty,+\infty)$，写成分段函数的形式

⑤ **求条件概率** 例 1 (3) 例 2 (3)

· 将条件（“|”之后的那个）代入条件密度函数，在指定范围内对“|”前的变量作积分即可

⑥ **独立性判断**

· 检查是否 $f_X(x)f_Y(y)=f(x,y)$ 或 $F_X(x)F_Y(y)=F(x,y)$

⑦ **求协方差** 例 1 (4)

· 由 $f(x,y)$ 计算 $E(XY)$，由 $f_X(x),f_Y(y)$ 计算 $E(X),E(Y)$，然后代入公式计算

⑧ **相关性判断** 例 1 (4)

· 若协方差大于 0 则正相关，小于 0 则负相关，等于 0 则不相关

2.历年考试典型例题

例 1 （19—20 春夏）设(X,Y)的联合密度函数

$$f(x,y)=\begin{cases}\frac{3y}{2},0<y<2x<2\\0,\quad 其他\end{cases}$$

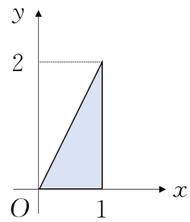
（1）求 $P(\max(X,Y)<1)$；

（2）分别求 X,Y 的边际密度函数 $f_X(x)$，$f_Y(y)$；

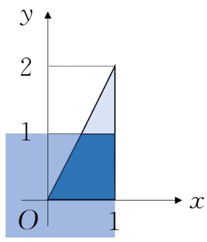
（3） $P(Y>0.5\mid X=0.5)$；

（4）判断 X 与 Y 是正相关，负相关，还是不相关？说明理由.

解 本题的非零区域如下图所示



（1） $P(\max(X,Y)<1)=P(X<1,Y<1)$，该区域与非零区域的交集为



对该区域进行积分：$P(\max(X,Y)<1)=\int_0^1 dy\int_{y/2}^1\frac{3y}{2}dx=\frac{1}{4}$

（2）① 求 $f_X(x)$：仅 $0<x<1$ 时，$f(x,y)$ 非 0，此时 $f_X(x)=\int_0^{2x}\frac{3y}{2}dy=3x^2$

因此 $f_X(x)=\begin{cases}3x^2,\quad 0<x<1\\0,\qquad 其他\end{cases}$

② 求 $f_Y(y)$：仅 $0<y<2$ 时，$f(x,y)$ 非 0，此时 $f_Y(y)=\int_{y/2}^1\frac{3y}{2}dx=\frac{3y}{2}(1-\frac{y}{2})$

因此 $f_Y(y)=\begin{cases}\frac{3y}{2}(1-\frac{y}{2}),\quad 0<y<2\\0,\qquad 其他\end{cases}$

（3）条件概率密度 $f_{Y|X}(y\mid x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}\frac{y}{2x^2},0<y<2x<2\\0,\qquad 其他\end{cases}$

$$\therefore\ P(Y>\frac{1}{2}|X=\frac{1}{2})=\int_{1/2}^{+\infty}f_{Y|X}(y|\frac{1}{2})\mathrm{d}y=\int_{1/2}^{1}2y\mathrm{d}y=\frac{3}{4}$$

（4）$E(XY)=\iint\limits_{D}xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{2x}xy\frac{3y}{2}\mathrm{d}y=\frac{4}{5}$

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

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

$$\therefore\ \mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{4}{5}-\frac{3}{4}\times 1=\frac{1}{20}>0$$

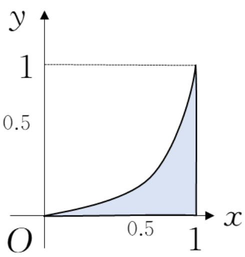
$\therefore$ X 与 Y 正相关

例 2 （17—18 **春夏 节选**）设(X,Y)的联合密度函数 $f(x,y)=\begin{cases}4x, & 0<y<x^2,0<x<1\\0, & 其他\end{cases}$：

（1）求(X,Y)的联合分布函数值$F(0.5,0.5)$；

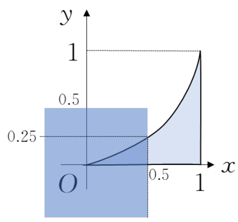
（2）分别求 X 与 Y 的边际概率密度函数 $f_X(x)$ 和 $f_Y(y)$，并判断 X 与 Y 是否相互独立；

解 本题的非零区域如下图所示：



（1）$F(0.5,0.5)=P(X\le 0.5,Y\le 0.5)$

$$=\int_0^{0.5}\mathrm{d}x\int_0^{x^2}4x\mathrm{d}y=\boxed{1/16}$$



（2）仅 $0<y<1$ 时，$f(x,y)$ 非 0，此时 $f_Y(y)=\int_{\sqrt{y}}^{1}4x\mathrm{d}x=2-2y$

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

仅 $0<x<1$ 时，$f(x,y)$ 非 0，此时 $f_X(x)=\int_0^{x^2}4x\mathrm{d}y=4x^3$

因此 $f_X(x)=\begin{cases}4x^3, & 0<x<1\\0, & 其他\end{cases}$

· 显然 $f_X(x)f_Y(y)\neq f(x,y)$，因此 X 与 Y 不独立