引言

二维联合分布全面地反映了二维随机变量(X,Y)的取值及其概率规律. 而单个随机变量X,Y也具有自己的概率分布. 那么要问:二者之间有什么关系呢?

这一节里,我们就来探求这个问题 .

第二节　边缘分布

一、边缘分布函数概念

二、离散型随机变量的边缘分布律

三、连续型随机变量的边缘密度函数

四、小节

3.2.1 边缘分布函数概念

1.定义

设$F(x,y)$为二维随机变量(X,Y)的分布函数，而X和Y又各自是一维随机变量，也具有分布函数，分别记为$F_X(x)$和$F_Y(y)$，分别称为(X,Y)关于X和关于Y的**边缘分布函数**（或**边际分布函数**）,简称为X和Y的**边缘分布函数**. (X,Y)的分布函数$F(x,y)$可唯一确定X和Y的边缘分布函数.

3.2.1 边缘分布函数概念

1.定义

$$F_X(x) = P(X \le x) = P(X \le x, Y < \infty) = \lim_{y \to +\infty} F(x, y) = F(x, \infty)$$

$$F_Y(y) = P(Y \le y) = P(X < x, Y \le y) = \lim_{x \to +\infty} F(x, y) = F(\infty, y)$$

一般地，若n维随机向量$(X_1, X_2, \dots, X_n)$的分布函数为$F(x_1, x_2, \dots, x_n)$，则X_i的边缘分布函数为

$$\boldsymbol{F_{X_i}(x_i) = F(+\infty, \dots, +\infty, x_i, +\infty, \dots, +\infty), i = 1, 2, \dots, n}$$

3.2.2 离散型随机变量的边缘分布律

2. 离散型随机变量的边缘分布律

设二维离散型随机变量(X,Y)的联合分布律为$P\{X=x_i,Y=y_j\}=p_{ij}$, $i,j=1,2,\dots$,记

$$P\{X=x_i\}=\sum_{j=1}^{\infty}P(X=x_i,Y=y_j)=\sum_{j=1}^{\infty}Pij=pi\cdot,\qquad \mathrm{i}=1,2,\dots,$$

$$P\{Y=y_j\}=\sum_{i=1}^{\infty}P(X=x_i,Y=y_j)=\sum_{i=1}^{\infty}Pij=p\cdot j,\qquad \mathrm{j}=1,2,\dots,$$

分别称 $pi\cdot$（$\mathrm{i}=1,2,\dots,$）和 $p\cdot j$（$\mathrm{j}=1,2,\dots,$）为关于X和关于Y的边缘分布律,简称X和Y的边缘分布律.

3.2.2 离散型随机变量的边缘分布律

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$ $P\{X=x_i\}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$\sum_j p_{1j}$ $p_{1\bullet}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$\sum_j p_{ij}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		
$p_{\cdot j}$	$\sum_i p_{i1}$	$\sum_i p_{i2}$		$\sum_i p_{ij}$		

$$P\{X=x_i\}=\sum_{j=1}^{+\infty} p_{ij},\ i=1,2,\cdots;$$

$$P\{Y=y_j\}=\sum_{i=1}^{+\infty} p_{ij},\ j=1,2,\cdots.$$

3.2.2 离散型随机变量的边缘分布律

则关于X的边缘分布律为

X	x_1	x_2	…	x_n	…
$P_{i\cdot}$	$p_{1\cdot}$	$p_{2\cdot}$	…	$p_{n\cdot}$	…

关于Y的边缘分布律为

Y	y_1	y_2	…	y_n	…
$P_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	…	$p_{\cdot n}$	…

二维离散型随机变量关于X和Y的边缘分布函数分别为

$$F_X(x) = F(x, +\infty) = \sum_{x_i \le x} \sum_{j=1}^{+\infty} p_{ij},$$

$$F_Y(y) = F(+\infty, y) = \sum_{y_j \le y} \sum_{i=1}^{+\infty} p_{ij}.$$

3.2.2 离散型随机变量的边缘分布律

例 设随机变量 X 在 1,2,3,4 四个整数中等可能的取值,另一个随机变量 Y 在 1~X 中等可能的取一整数值.试求 (X,Y) 的分布律,同时求出关于 X,Y 的边缘分布律.

解: $\{X=i, Y=j\}$ 的取值情况是: $i=1, 2, 3, 4$, j 取不大于 i 的正整数.

由乘法公式得: $P\{X=i, Y=j\}=P\{Y=j \mid X=i\}P\{X=i\}=\frac{1}{i}\cdot\frac{1}{4}$, $i=1, 2, 3, 4, j\leqslant i$.

(X, Y) 的分布律为

$X \backslash Y$	1	2	3	4
1	$\frac{1}{4}$	0	0	0
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

$X \backslash Y$	1	2	3	4	$p_{i\cdot}$
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$	

3.2.2 离散型随机变量的边缘分布律

联合分布律为

$X \backslash Y$	1	2	3	4	$p_{i\cdot}$
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$	

则关于X的边缘分布律为

X	1	2	3	4
$p_{i\cdot}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

关于Y的边缘分布律为

Y	1	2	3	4
$p_{\cdot j}$	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$

注意 **联合分布** $\overset{\longrightarrow}{\not\longleftarrow}$ **边缘分布**

3.2.3 连续型随机变量的边缘概率密度

3. 连续型随机变量的边缘密度函数

定义 设二维连续型随机变量(X,Y)具有密度函数$f(x,y)$,其分布函数为

$F(x,y)$,

$$F_X(x)=P(X\le x,Y<\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(x,y)dy\right]dx$$

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

为随机变量(X,Y)关于X的边缘密度函数(简称X的边缘密度函数).

同理可得随机变量(X,Y)关于Y的边缘密度函数（简称Y的边缘密度函数）

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

例3.2.3 设二维随机变量(X,Y)的密度函数$f(x,y)$为

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

试求X和Y的边缘密度函数

解 X的边缘密度函数为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_{-x}^{x}dy, & 0<x<1,\\ 0, & \text{其他}.\end{cases}=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

Y的边缘密度函数为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}\int_{-y}^{1}dx, & -1<y\le 0,\\ \int_{y}^{1}dx, & 0<y<1,\\ 0, & \text{其他}.\end{cases}=\begin{cases}1+y, & -1<y\le 0,\\ 1-y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

3.2.3 连续型随机变量的边缘概率密度

练习 设二维随机变量(X,Y)具有联合密度，

$$f(x,y)=\begin{cases}6, & x^2 \le y \le x,\\ 0, & \text{其他}\end{cases}$$

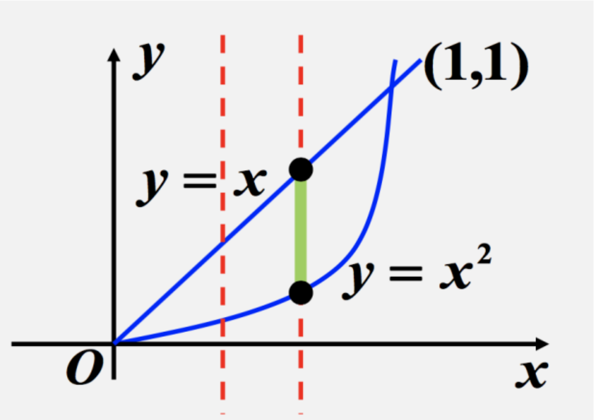
试求X和Y的边缘密度函数$f_X(x), f_Y(y)$

解 X的边缘密度函数为

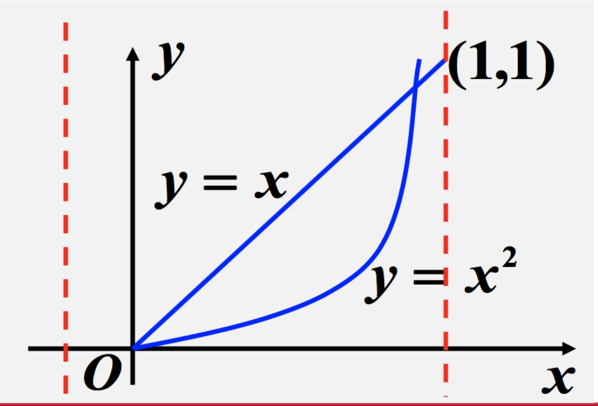
$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$$

$$=\begin{cases}\int_{x^2}^{x} 6dy, & 0 \le x \le 1,\\ 0, & \text{其他.}\end{cases}$$

$$f_X(x) = \begin{cases}6(x - x^2, & 0 \le x \le 1,\\ 0, & \text{其他.}\end{cases}$$



$0 \le x \le 1$



$x < 0$或$x > 1$

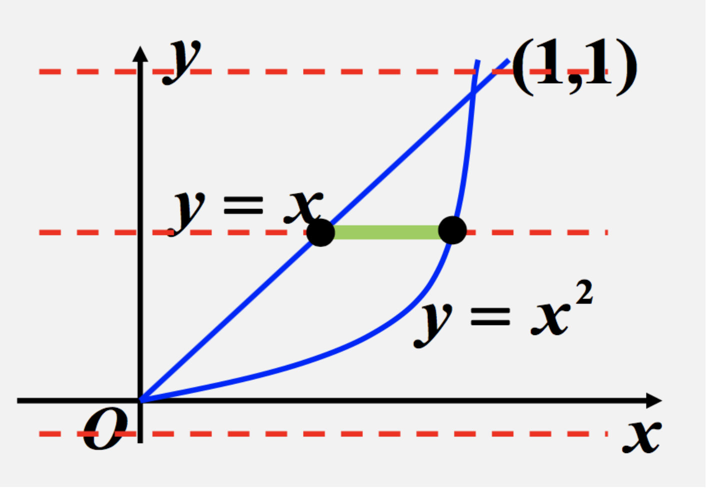
练习 设二维随机变量(X,Y)具有联合密度,

$$f(x,y)=\begin{cases}6, & x^2\le y\le x,\\ 0, & \text{其他}\end{cases}$$ 试求X和Y的边缘密度函数$f_X(x), f_Y(y)$

解 Y的边缘密度函数为

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

$$=\begin{cases}\displaystyle\int_y^{\sqrt{y}}6dx & 0\le y\le 1,\\ 0, & \text{其他.}\end{cases}$$



$0\le y\le 1$

$$f_Y(y)=\begin{cases}6(\sqrt{y}-y), & 0\le y\le 1,\\ 0, & \text{其他.}\end{cases}$$

3.2.3 连续型随机变量的边缘概率密度

例3.2.4 **设(X,Y)~N($\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$)，即(X,Y)具有概率密度，**

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

$$-\infty<x<+\infty,-\infty<y<+\infty$$

试求X和Y的边缘密度函数$f_X(x)$，$f_Y(y)$.

解:由于

$$\frac{(y-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}=\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2-\rho^2\frac{(x-\mu_1)^2}{\sigma_1^2}$$

3.2.3 连续型随机变量的边缘概率密度

$$f_X(x)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{\frac{-1}{2\left(1-\rho^2\right)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\rho^2\frac{(x-\mu_1)^2}{\sigma^2}\right]_{\sigma_1^2}}$$

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

令: $t=\dfrac{1}{\sqrt{1-\rho^2}}\left(\dfrac{y-\mu_2}{\sigma_2}-\rho\dfrac{x-\mu_1}{\sigma_1}\right)$ 则: $dt=\dfrac{1}{\sigma_2\sqrt{1-\rho^2}}dy$

$$f_X(x)=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt$$

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

$X\sim N(\mu_1,\sigma_1^2)$

3.2.3 连续型随机变量的边缘概率密度

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

$Y\sim N(\mu_2,\sigma_2^2)$.

即X的边缘分布是正态分布，且$X\sim N(\mu_1,\sigma_1^2)$，Y的边缘分布也是正态分布，$Y\sim N(\mu_2,\sigma_2^2)$.

说明二维正态分布的两个边缘分布都是一维正态分布，且不依赖参数ρ. 对于给定的参数μ_1, μ_2,σ_1^2, σ_2^2，不同的参数ρ对应不同的二维正态分布，但边缘分布一样，说明边缘分布一般确定不了联合分布。

联合分布 $\underset{\not\longleftarrow}{\longrightarrow}$ **边缘分布**

小结

1. 与一维情形相对照，介绍了二维随机变量的边缘分布.

$$F_X(x) = P(X \le x) = P(X \le x, Y < \infty) = \lim_{y \to +\infty} F(x, y) = F(x, \infty)$$

$$F_Y(y) = P(Y \le y) = P(X < x, Y \le y) = \lim_{x \to +\infty} F(x, y) = F(\infty, y)$$

2. 请注意联合分布和边缘分布的关系:

由联合分布可以确定边缘分布;

但由边缘分布一般不能确定联合分布.