概率论与数理统计

Probability and Statistics

—— 概率论与数理统计教学组——

哈尔滨工程大学

第3章 多维随机变量及其分布

3.2 边缘分布

学习要点

- 二维离散型随机变量的边缘分布
- 二维连续型随机变量的边缘分布

一、 边缘分布引言

若(X,Y)为二维随机变量，分布函数为$F(x,y)$.

分量X为一维随机变量，分布函数$F_X(x)$称为(X,Y)的关于X的**边缘分布函数**；

分量Y为一维随机变量，分布函数$F_Y(y)$称为(X,Y)的关于Y的**边缘分布函数**；

$$F_X(x) = P\{X \le x\} = P\{(X \le x) \cap (Y \le +\infty)\}$$

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

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

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

二、二维离散型随机变量的边缘分布

若二维离散型随机变量(X,Y)具有分布函数$F(x,y)$，则分量X,Y都是一维离散型随机变量， X,Y的**边缘分布函数**分别为:

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

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

边缘分布

(X,Y)关于 X 的分布律为

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

同理(X,Y)关于Y 的分布律为

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

称 $p_{i\cdot}\ (i=1,2,\cdots)$和 $p_{\cdot j}\ (j=1,2,\cdots)$为$(X,Y)$关于$X$ 和Y 的**边缘分布律**.

边缘分布

例 1 设袋中装有 2 个白球、3 个红球，现从袋中无放回地随机抽取两次，定义随机变量 X,Y 如下：

$$X=\begin{cases}1 & \text{第一次取出白球} \\ 0 & \text{第一次取出红球}\end{cases},\quad Y=\begin{cases}1 & \text{第二次取出白球} \\ 0 & \text{第二次取出红球}\end{cases},$$

求随机变量 (X,Y) 的联合分布律及边缘分布律.

解 由题意可知

$Y \backslash X$	0	1	$p_{\cdot j}$
0	$3/5\times2/4$	$+\ 2/5\times3/4$	$=3/5$
1	$3/5\times2/4$	$+\ 2/5\times1/4$	$=2/5$
$p_{i\cdot}$	$=3/5$	$=2/5$	

表格“边缘”

例 2 若(X,Y)的联合分布律为

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

若(X,Y)的联合分布律为

$Y \backslash X$	0	1
0	0.2	0.3
1	0.3	0.2

由(X,Y)的联合分布，

可以确定边缘X和Y的分布；

则其边缘分布律为

X	0	1
p_k	0.5	0.5

Y	0	1
p_k	0.5	0.5

由边缘X和Y的分布，

不能确定(X,Y)的联合分布。

三、二维连续型随机变量的边缘分布

对于二维连续型随机变量(X,Y)，设它的联合概率密度为 $f(x,y)$，由

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}[\int_{-\infty}^{+\infty}f(x,y)dy]dx$$

因此

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

边缘概率密度函数

同理

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

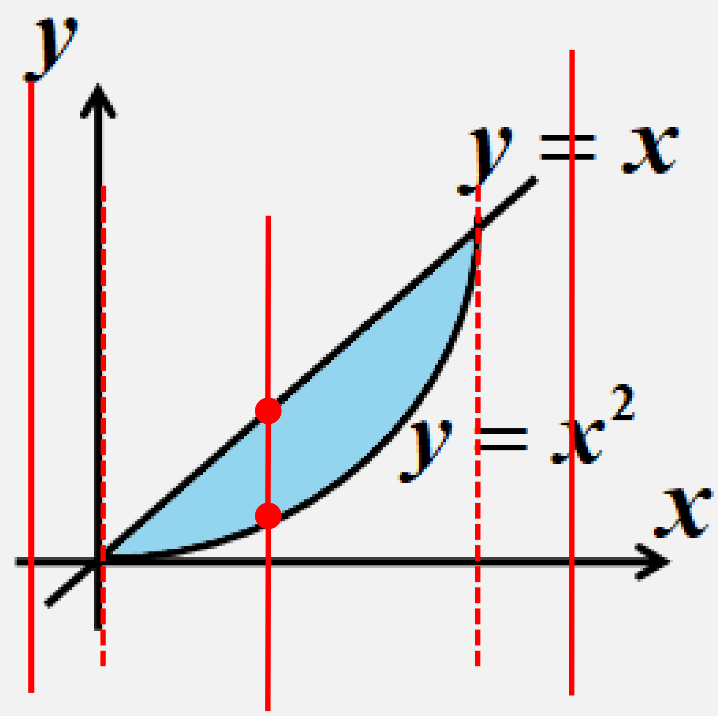
例 3 设二维随机变量(X,Y)的联合概率密度为

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

求边缘概率密度 $f_X(x), f_Y(y)$.

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

$$=\begin{cases}\int_{x^2}^{x} 6dy=6\left(x-x^2\right), & 0\le x\le 1\\ 0, & \text{其他}\end{cases}$$



当$0\le x\le 1$时 当$x\notin[0,1]$时

例 3 设二维随机变量(X,Y)的联合概率密度为

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

求边缘概率密度 $f_X(x), f_Y(y)$.

解 即

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{x^2}^{x}6\mathrm{d}y=6\left(x-x^2\right), & 0\le x\le 1,\\ 0, & \text{其他}.\end{cases}$$

同理 $$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{y}^{\sqrt{y}}6\mathrm{d}x=6\left(\sqrt{y}-y\right), & 0\le y\le 1,\\ 0, & \text{其他}.\end{cases}$$

四、二维正态分布

设二维随机变量(X,Y)在$-\infty<x<+\infty,\ -\infty<y<+\infty$上概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot$$

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

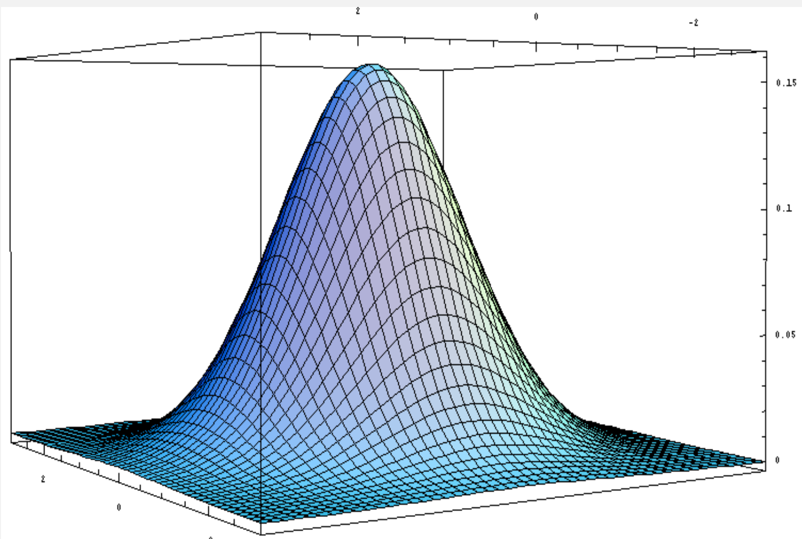
其中$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 都是常数，且 $\sigma_1>0,\ \sigma_2>0,\ -1<\rho<1,$

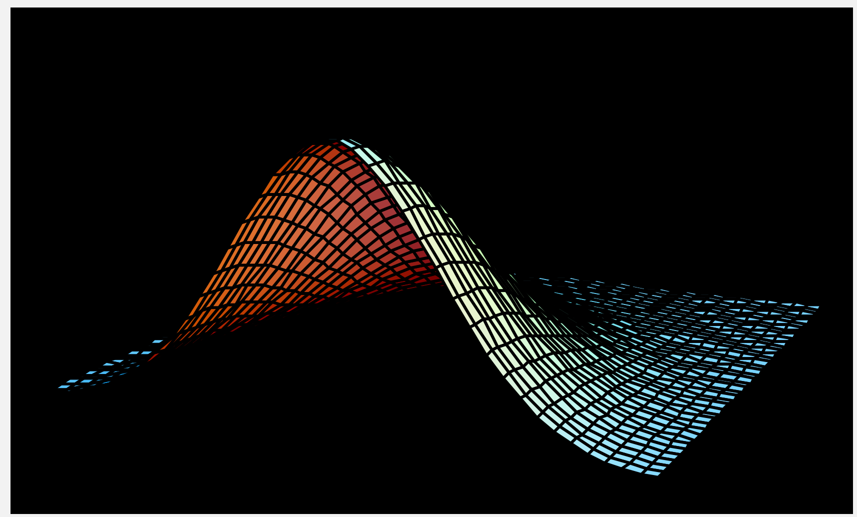
称(X,Y)服从参数为$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$的**二维正态分布**,

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

边缘分布

二维正态分布密度函数图形及剖面图





例 4 设(X,Y)服从二维正态分布，即$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求$X$和$Y$的边缘概率分布.

解 (X,Y)在$-\infty<x<+\infty$，$-\infty<y<+\infty$上概率密度为

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

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot$$

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

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

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\left\{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right\}\cdot$$

$$\exp\left\{-\frac{1}{2}\cdot\frac{1}{(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2\right\}$$

解 (X,Y)在$-\infty<x<+\infty,\ -\infty<y<+\infty$上概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\left\{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right\}\cdot$$

$$\exp\left\{-\frac{1}{2}\cdot\frac{1}{(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2\right\}$$

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

$$=\frac{1}{2\pi\sigma_1}\cdot\exp\left\{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right\}\cdot$$

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

$$f_X(x)=\frac{1}{2\pi\sigma_1}\cdot e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}}\cdot$$

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

$$=\frac{1}{2\pi\sigma_1}\cdot e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}}\cdot\int_{-\infty}^{+\infty}e^{-\frac{1}{2}t^2}\,\mathrm{d}t$$

$$\Phi(+\infty)=1$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}\cdot e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}}\cdot\overbrace{\boxed{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}\,\mathrm{d}t}}^{\Phi(+\infty)=1}=\frac{1}{\sqrt{2\pi}\sigma_1}\cdot e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}}$$

因此，边缘$X\sim N\left(\mu_1,\sigma_1^2\right)$**，同理，边缘**$Y\sim N\left(\mu_2,\sigma_2^2\right)$.

注意： 在例 4 中看出，当$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$时，

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot$$

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

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

$\mu_1,\mu_2,\sigma_1,\sigma_2$固定不变，当$\rho$变化时，边缘$X$和$Y$的分布不变，而$(X,Y)$的联合分布发生变化。

由(X,Y)的联合分布，可以确定边缘X和Y的分布；
由边缘X和Y的分布，**不能**确定(X,Y)的联合分布。

小结 边缘分布

- 二维**离散型**随机变量的边缘分布
- 二维**连续型**随机变量的边缘分布

谢谢您的观看

HARBIN ENGINEERING UNIVERSITY