2.3 二元随机变量

定义 2.5 如果每次试验的结果对应着一组确的实数$(\xi_1,\xi_2,\ldots,\xi_n)$, 它们是随试验结果不同而变化的$n$个随机变量, 并且对任何一组实数$x_1,x_2,\ldots,x_n$, 事件"$\xi_1\le x_1,\xi_2\le x_2,\ldots,\xi_n\le x_n$" 有确定的概率, 则称$n$个随机变量的整体$(\xi_1,\xi_2,\ldots,\xi_n)$为一个$n$元随机变量(或$n$元随机向量)

定义 2.6 称n元函数

$$F(x_1,x_2,\ldots,x_n)=P(\xi_1\le x_1,\xi_2\le x_2,\ldots,\xi_n\le x_n)$$

$$(x_1,x_2,\ldots,x_n)\in R^n$$

为n元随机变量的分布函数.

注意

事件"$\xi_1 \le x_1, \xi_2 \le x_2, \ldots, \xi_n \le x_n$"表示$n$个事件 $\{\xi_1 \le x_1\}, \{\xi_2 \le x_2\}, \ldots, \{\xi_n \le x_n\}$的交事件, 即

$$\{\xi_1 \le x_1\} \cap \{\xi_2 \le x_2\} \cap \ldots \cap \{\xi_n \le x_n\}$$

如前所述, n个事件的交事件通常不好计算, 要利用乘法法则来进行计算. 即利用公式

$$P(A_1 A_2 \ldots A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \ldots$$

$$\ldots P(A_n | A_1 A_2 \ldots A_{n-1})$$

(一)离散型

1. 联合分布

定义 2.7 如果二元随机变量(ξ,η)所有可能取的数对为有限或可列个, 并且以确定的概率取各个不同的数对, 则称(ξ,η)为二元离散型随机变量.

为了直观, 可以把(ξ,η)所有的可能取值及相应概率列成表, 称为(ξ,η)的联合概率分布表

$\xi \backslash \eta$	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

也可以用一系列等式来表示二元离散型随机变量(ξ,η)的联合概率分布.

$$P\{\xi=x_i,\eta=y_j\}=p_{ij} \qquad (i,j=1,2,\dots)$$

这都被称作ξ与η的联合分布律, 具有性质:

$$(1)\quad p_{ij} \geq 0$$

$$(2)\quad \sum_i \sum_j p_{ij} = 1$$

例1 同一品种的5个产品中, 有2个正品, 每次从中取1个检验质量, 不放回地抽取, 连续2次, 记"$\xi_k=0$"表示第k次取到正品, 而"$\xi_k=1$"为第k次取到次品(k=1,2). 写出(ξ_1,ξ_2)的联合分布律.

解 按乘法公式有

$$P\{\xi_1=0,\xi_2=0\}=P\{\xi_1=0\}P\{\xi_2=0\mid\xi_1=0\}=\frac{2}{5}\times\frac{1}{4}=0.1$$

$$P\{\xi_1=0,\xi_2=1\}=\frac{2}{5}\times\frac{3}{4}=0.3$$

$$P\{\xi_1=1,\xi_2=0\}=\frac{3}{5}\times\frac{2}{4}=0.3$$

$$P\{\xi_1=1,\xi_2=1\}=\frac{3}{5}\times\frac{2}{4}=0.3$$

列成概率分布表为

$\xi_2 \backslash \xi_1$	0	1
0	0.1	0.3
1	0.3	0.3

边缘分布与联合分布的关系

二元随机变量(ξ,η)中, 分量ξ(或η)的概率分布称为(ξ,η)的关于ξ(或η)的边缘分布. 如果已知(ξ,η)的联合分布为

$$P\{\xi=x_i,\eta=y_j\}=p_{ij} \quad (i, j=1,2,\dots)$$

则

$$P(\xi = x_i) = \sum_j P(\xi = x_i, \eta = y_j) = \sum_j p_{ij} = p_i^{(1)}$$

$$(i = 1,2,\dots)$$

$$P(\xi = y_j) = \sum_i P(\xi = x_i, \eta = y_j) = \sum_i p_{ij} = p_j^{(2)}$$

$$(j = 1,2,\dots)$$

例2

将两封信随机地往编号为1,2,3,4的4个邮筒内投. ξ_i表示第i个邮筒内信的数目(i=1,2). 写出(ξ_1,ξ_2)的联合分布及(ξ_1,ξ_2)中关于ξ_1的边缘分布

解 试验共有4^2=16种不同的等可能结果

$$p_{00} = P\{\xi_1 = 0, \xi_2 = 0\} = \frac{4}{16}$$

$$p_{01} = P\{\xi_1 = 0, \xi_2 = 1\} = \frac{2\times 2}{16} = \frac{4}{16}$$

$$p_{10} = p_{01} = \frac{4}{16} \quad p_{11} = \frac{2}{16}$$

$$p_{02} = p_{20} = \frac{1}{16} \quad p_{12} = p_{21} = p_{22} = 0$$

计算结果列于下表并计算ξ_1的边缘分布

$\xi_1 \backslash \xi_2$	0	1	2	$p_i^{(1)}$
0	4/16	4/16	1/16	9/16
1	4/16	2/16	0	6/16
2	1/16	0	0	1/16

上表计算出的ξ_1的边缘分布可列成下表

ξ_1	0	1	2
P	9/16	6/16	1/16

条件分布 对于二元离散型随机变量(ξ,η), 如果$P\{\eta=y_j\}>0$, 称$p_{ij}/p_j^{(2)}(i=1,2,\ldots)$为在$\eta=y_j$条件下关于$\xi$的条件分布, 记为

$$P\{\xi = x_i \mid \eta = y_j\} = \frac{p_{ij}}{p_j^{(2)}} \quad (i = 1,2,\ldots)$$

显然$P\{\xi=x_i|\eta=y_j\}$是非负的, 并且对于所有的i, 它们的和为1, 同样地, 若$p_i^{(1)}>0$, 称

$$P\{\eta = y_j \mid \xi = x_i\} = \frac{p_{ij}}{p_i^{(1)}} \quad (j = 1,2,\ldots)$$

为在$\xi=x_i$条件下关于η的分布.

连续型 二元连续型随机变量是用联合概率密度函数$\varphi(x,y)$来描述的, 它具有性质

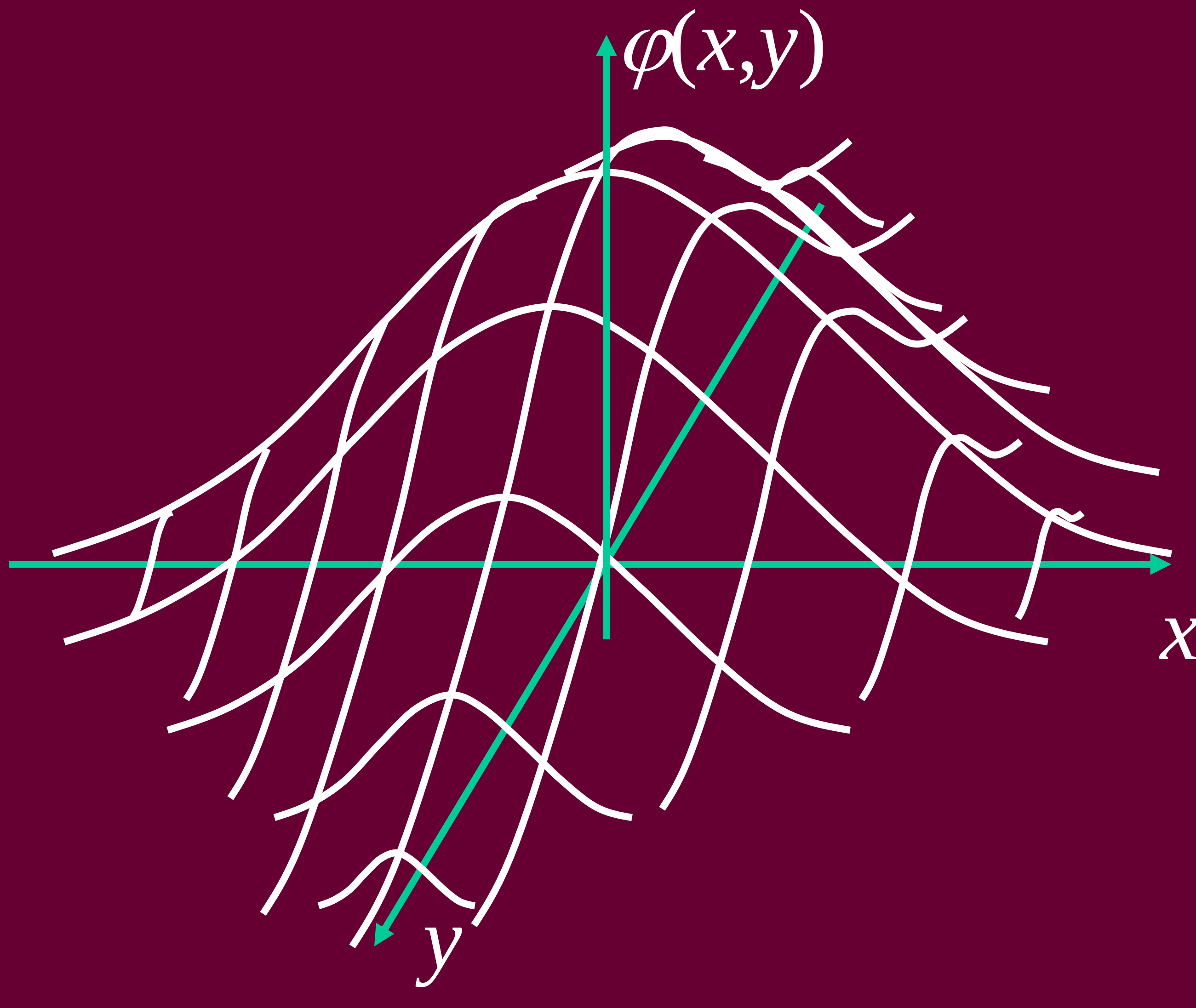
(1) 对任意实数x, y, $\varphi(x, y) \geq 0$

$$(2) \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(x, y)dxdy = 1$$

因此对于平面上任何可积区域S, (ξ, η)落在此区域内的概率是$\varphi(x,y)$在S上的二重积分, 即

$$P\{(\xi, \mu) \in S\} = \iint_S \varphi(x, y)dxdy$$

二元概率密度函数$\varphi(x,y)$从图形上看是在xoy平面上方的一个曲面, 包围着下方的体积为1.



显然, 对任意实数$a<b$及$c<d$, 有

$$P\{a<\xi\le b, c<\eta\le d\}=\int_a^b\int_c^d \varphi(x,y)dydx$$

(ξ,η)的分布函数$F(x,y)$也可由下式求出:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}\varphi(s,t)dtds$$

(ξ,η)关于ξ及η的边缘分布函数可按下式求出

$$F_\xi(x) = P\{\xi \le x\} = P\{\xi \le x, -\infty < \eta < +\infty\}$$

$$= \int_{-\infty}^{x} ds \int_{-\infty}^{+\infty} \varphi(s,t)\,dt$$

$$F_\eta(y) = P\{\eta \le y\} = P\{-\infty < x < +\infty, \eta \le y\}$$

$$= \int_{-\infty}^{y} dt \int_{-\infty}^{+\infty} \varphi(s,t)\,ds$$

若记

$$\varphi_1(x)=\varphi_\xi(x)=\int_{-\infty}^{+\infty}\varphi(x,y)dy$$

$$则 F_\eta(x)=\int_{-\infty}^{x}\varphi_1(s)ds$$

称$\varphi_1(x)$或φ_ξ(y)是(ξ,η)中关于ξ的边缘概率密度. 同样地记

$$\varphi_2(y)=\varphi_\eta(y)=\int_{-\infty}^{+\infty}\varphi(x,y)dx$$

则称φ_2(y)或φ_η(y)是(ξ,η)中关于η的边缘概率密度.

若 $\varphi_2(y)>0$, 称

$$\varphi(x\,|\,y)=\frac{\varphi(x,y)}{\varphi_2(y)}$$

为在 $\eta=y$ 条件下, 关于 ξ 的条件概率密度

称

$$\varphi(y\,|\,x)=\frac{\varphi(x,y)}{\varphi_1(x)}$$

为在 $\xi=x$ 条件下, 关于 η 的条件概率密度

随机变量的独立性

两个随机变量ξ和η是相互独立的, 是指的其中一个变量取任意值的事件和另一个变量取任意值的事件总是相互独立的.

严格的定义为:

定义 2.9 对于任何实数x,y, 如果二元随机变量(ξ,η)的联合分布函数$F(x,y)$等于ξ和η的边缘分布函数的乘积, 即

$$F(x,y)=F_\xi(x)F_\eta(y)$$

则称随机变量ξ与η相互独立.

离散型 ξ与η相互独立的充要条件是对一切$i,j=1,2,\ldots$

$$p_{ij}=p_i^{(1)}p_j^{(2)}$$

在给定离散型随机变量的概率分布表的情况下，如果要判定其不独立往往容易, 只要任找一个p_{ij}不等于边缘概率$p_i^{(1)}$和$p_j^{(2)}$的乘积就可断定其不独立. 经常的快捷办法就是, 只要发现联合概率分布表中有0存在, 就基本可以认为这两个随机变量不独立了.

而如果要判定其独立, 则需要验证每一个p_{ij}是否为各个边缘概率的乘积.

例5 本节例2的两个随机变量ξ_1和ξ_2是否相互独立?

ξ_1 \ ξ_2	0	1	2
0	4/16	4/16	1/16
1	4/16	2/16	0
2	1/16	0	0

解 $p_{22}=0\neq p_2^{(1)}p_2^{(2)}=(1/16)(1/16)$

因此ξ_1和ξ_2不独立.

连续型

如ξ和η为连续型随机变量, 则它们相互独立的充分必要条件为, 对任何实数x, y

$$\varphi(x, y)=\varphi_1(x)\varphi_2(y)=\varphi_\xi(x)\varphi_\eta(y)$$

当一个二元函数$f(x, y)$可写成两个单变量的函数乘积$f(x, y)=g(x)h(y)$时, 称其为可分离变量的. 不难证明如果ξ和η的联合概率密度$\varphi(x,y)$可分离变量的, 它们就是相互独立的, 反之亦然.

例6 两个随机变量x1与x2相互独立, 其概率密度为

$$\varphi_i(x_i)=\frac{1}{\sqrt{2\pi}\sigma_i}e^{-\frac{1}{2}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2}\qquad (i=1,2)$$

求它们的联合概率密度.

解:

$$\varphi(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

例3.2.4 已知 (X, Y) 的联合密度为

$$p(x,y)=\begin{cases} e^{-x-y}, & x>0,\ y>0; \\ 0, & 其\ 他. \end{cases}$$

问 X 与 Y 是否独立？

解: 边际分布密度分别为:

$$p(x)=\begin{cases} \int_0^{+\infty} e^{-(x+y)}\mathrm{d}y=e^{-x} & x>0 \\ 0 & x\le 0 \end{cases} \qquad p(y)=\begin{cases} e^{-y}, & y>0 \\ 0, & y\le 0 \end{cases}$$

所以 X 与 Y 独立。 注意：$p(x, y)$ 可分离变量.

例4 设(X,Y)的概率密度是

$$f(x,y)=\begin{cases} cy(2-x), & 0\le x\le 1,\quad 0\le y\le x \\ 0, & \text{其它} \end{cases},$$

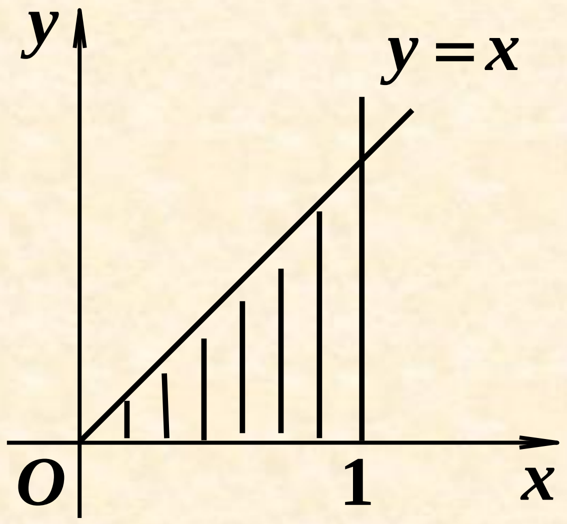
求 (1) c 的值; (2) 两个边缘密度.

解 (1) 由 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy=1$ 确定 c.

$$\int_0^1\left[\int_0^x cy(2-x)dy\right]dx$$

$$=c\int_0^1[x^2(2-x)/2]dx$$

$$=5c/24=1 \longrightarrow c=24/5.$$



例4 设(X,Y)的概率密度是

$$f(x,y)=\begin{cases} cy(2-x), & 0\le x\le 1,\quad 0\le y\le x \\ 0, & \text{其它} \end{cases},$$

求 (1) c 的值;　(2) 两个边缘密度.

解 (2) $f_X(x)=\int_0^x \frac{24}{5}y(2-x)dy=\frac{12}{5}x^2(2-x),$

$$0\le x\le 1$$

$$f_Y(y)=\int_y^1 \frac{24}{5}y(2-x)dx=\frac{24}{5}y\left(\frac{3}{2}-2y+\frac{y^2}{2}\right),$$

$$0\le y\le 1$$

知识导航 结束放映

例4 设(X,Y)的概率密度是

$$f(x,y)=\begin{cases} cy(2-x), & 0\le x\le 1,\quad 0\le y\le x \\ 0, & \text{其它} \end{cases},$$

求 (1) c 的值; (2) 两个边缘密度.

解 (2) 即

$$f_X(x)=\begin{cases} \frac{12}{5}x^2(2-x), & 0\le x\le 1 \\ 0, & \text{其它} \end{cases}$$

$$f_Y(y)=\begin{cases} \frac{24}{5}y\left(\frac{3}{2}-2y+\frac{y^2}{2}\right), & 0\le y\le 1 \\ 0, & \text{其它} \end{cases}.$$

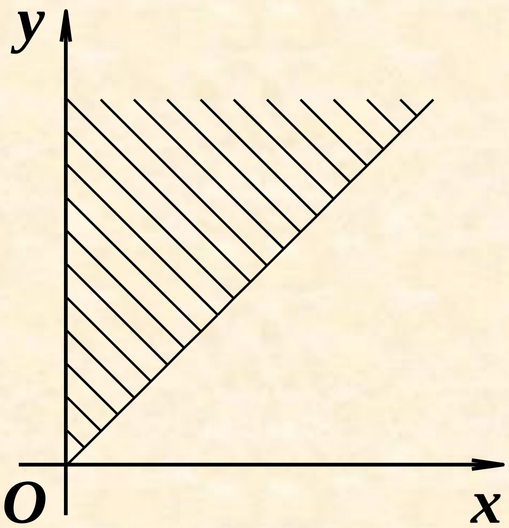
完

例6 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0, & \text{其它} \end{cases},$$

(1) 求 X与Y 的边缘概率密度，并判断X 与Y 是否相互独立;

例6 (1) 求 X与 Y的边缘概率密度，并判断 X 与Y 是否相互独立;



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

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

当 $x \le 0$ 时，$f_X(x)=0,$

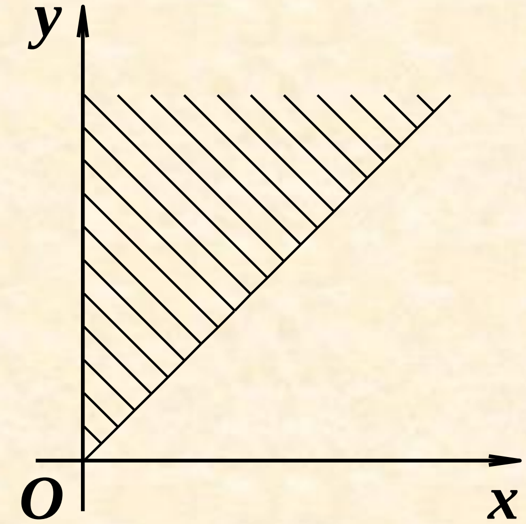
当 $x>0$ 时，$f_X(x)=\int_x^{+\infty} e^{-y}dy = e^{-x},$

所以 $f_X(x)=\begin{cases} e^{-x}, & x>0 \\ 0, & x\le 0 \end{cases}$，类似可得

$$f_Y(y)=\begin{cases} ye^{-y}, & y>0 \\ 0, & y\le 0 \end{cases}.$$

知识导航 结束放映

例6 (1) 求 X与 Y的边缘概率密度，并判断 X 与Y是否相互独立；



解 所以 $f_X(x)=\begin{cases} e^{-x}, & x>0 \\ 0, & x\le 0 \end{cases}$,

类似可得 $f_Y(y)=\begin{cases} ye^{-y}, & y>0 \\ 0, & y\le 0 \end{cases}$.

由于当 $0<x<y$ 时

$$f_X(x)\cdot f_Y(y)\neq f(x,y)$$

故 X 与Y不相互独立.