Ch 5 多维随机变量及其分布

回顾前一次课

正态分布 $X \sim N(\mu, \sigma^2)$ $\mathrm{E}(X) = \mu, \quad \mathrm{Var}(X) = \sigma^2$

$X \sim N(0,1)$的分布函数

$$P(X \geq \epsilon) \leq \frac{1}{2} e^{-\frac{\epsilon^2}{2}} \qquad P(|X| \geq \epsilon) \leq \; min\left(1, \sqrt{\frac{2}{\pi}} \frac{e^{-\frac{\epsilon^2}{2}}}{\epsilon}\right)$$

已知连续随机变量X的概率密度为$f_X(x)$，求随机变量$Y = g(X)$的概率密度$f_Y(y)$？

- 求解Y=g(X)的分布函数$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{g(x) \leq y} f_X(x) dx$
- 利用分布函数和概率密度关系求解密度函数$f_Y(y) = F_Y'(y)$

回顾前一次课

随机变量X的概率密度为$f_X(x)$，其中$x \in (-\infty, +\infty)$．函数$y = g(x)$处处可导且严格单调（即$g'(x) > 0$ 或 $g'(x) < 0$)，令其反函数$x = g^{-1}(y) = h(y)$，则随机变量$Y = g(X)$ 的概率密度为

$$f_Y(y) = \begin{cases} f_X\big(h(y)\big)|h'(y)| & y \in (\alpha, \beta) \\ 0 & \text{其它} \end{cases}$$

$\alpha = \min\{g(-\infty), g(+\infty)\}$，$\beta = \max\{g(-\infty), g(+\infty)\}$

二维随机变量

很多随机现象往往由两个或多个随机因素造成的，需用多个随机变量描述．如导弹攻击点的坐标 (经度、纬度)，学生的高考成绩 (语文、数学、英语等).

设$X = X(\omega)$和$Y = Y(\omega)$为定义在样本空间Ω上的随机变量, 由它们构成的向量(X, Y)称为**二维随机向量**

二维随机向量又称二维随机变量，需将(X, Y)看作一个整体, 不能分开, 在几何上(X, Y)可看作平面上的随机点

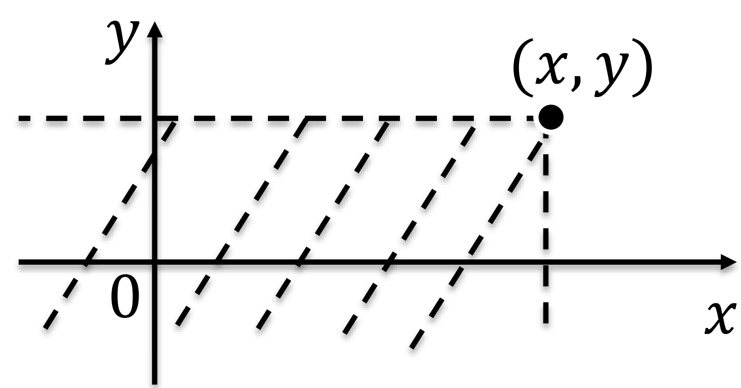
二维随机变量的分布函数

设(X,Y)为二维随机变量，对任意实数$x\in(-\infty,+\infty)$和$y\in(-\infty,+\infty)$，称

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

为**二维随机变量(X,Y)的分布函数**，或称为**随机变量X和Y的联合分布函数**

分布函数$F(x,y)$几何意义:
随机点(X,Y) 落入以(x,y)为右上定点无穷矩形的概率



二维随机变量分布函数性质

- 分布函数$F(x,y)$对每个变量单调不减
  - 固定y, 当$x_1 > x_2$时有$F(x_1,y) \geq F(x_2,y)$
  - 固定x, 当$y_1 > y_2$时有$F(x,y_1) \geq F(x,y_2)$
- 对任意$x \in (-\infty,+\infty)$和$y \in (-\infty,+\infty)$，分布函数$F(x,y) \in [0,1]$, 且

$$F(+\infty,+\infty) = 1$$

$$F(-\infty,y) = F(x,-\infty) = F(-\infty,-\infty) = 0$$

- 分布函数$F(x,y)$关于每个变量右连续

概率估计

根据分布函数可推导概率:

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2)$$
$$= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

随机变量(X, Y)的联合分布函数为$F(x, y)$，如果将随机变量X和Y分别看, 依然为随机变量

如何从联合分布函数$F(x, y)$研究随机变量X和Y的分布函数$F_X(x)$和$F_Y(y)$

边缘分布函数

设二维随机变量(X,Y)的联合分布函数为$F(x,y)$, 称

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

为随机变量X的边缘分布函数.

同理定义随机变量Y的边缘分布函数为:

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

例题

设二维随机变量(X,Y)的联合分布函数为

$$F(x,y)=A\left(B+\arctan\frac{x}{2}\right)\left(C+\arctan\frac{y}{3}\right)$$

求随机变量X与Y的边缘分布函数和概率$P(Y>3)$

随机变量的独立性

随机事件的独立性：$P(AB) = P(A)P(B)$

设X, Y为二维随机变量，对任意$x, y \in (-\infty, +\infty)$，若事件$X \le x$和$Y \le y$相互独立, 即

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$$

等价于 $F(x, y) = F_X(x)\, F_Y(y)$

则称随机变量***X*与*Y*相互独立**

定理： 设随机变量X与Y相互独立，则$f(X)$与$g(Y)$相互独立（其中$f(x)$和$g(y)$是连续或分段连续函数）

二维离散型随机变量

若二维随机变量 (X,Y) 的取值是有限个或无限可列的，称 (X,Y) 为二维离散型随机变量

设离散型随机变量(X,Y)的取值分别为(x_i,y_j)，$i,j=1,2,\cdots$，则称

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

为**(X,Y)的联合分布列**

$X \backslash Y$	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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$

性质：$p_{ij}\geq 0$和$\sum_{i,j}p_{ij}=1$

边缘分布列

根据二维随机变量(X,Y)的联合分布列p_{ij}

随机变量X的边缘分布列

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

随机变量Y的边缘分布列

$$P(Y=y_i)=\sum_{i=1}^{+\infty}P(X=x_i,Y=y_j)=\sum_{i=1}^{+\infty}p_{ij}=p_{\cdot j}$$

边缘分布列

二维随机变量联合和边缘分布表示在同一个表格

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

例题

有三个数1,2,3，随机变量X表示从这三个数中随机地抽取一个数，随机变量Y表示从1到X中随机抽取一个数. 求(X, Y)的联合分布列和边缘分布列

离散随机变量的独立性

对离散型随机变量(X, Y), 若对所有(x_i, y_j)有

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

即$p_{ij} = p_{i\cdot}\, p_{\cdot j}$，称**离散随机变量X与Y相互独立**

定理：对离散型随机变量(X, Y), 以下两种定义等价

$$p_{ij} = p_{i\cdot}\, p_{\cdot j} \leftrightarrow F(x_i, y_j) = F_X(x_i)\, F_Y(y_j)$$

离散随机变量的独立性

定理： 设离散随机变量X和Y独立，则对任意集合$A \in \mathrm{R}$, $B \in \mathrm{R}$, 有事件$X \in A$和$Y \in B$独立

例题： 设离散型随机变量X, Y独立，求解(X, Y)的联合分布列

$X \backslash Y$	y_1	y_2	y_3	$p_{i\cdot}$
x_1		$1/8$		
x_2	$1/8$			
$p_{\cdot j}$	$1/6$			

例题

将两个球A, B放入编号为1,2,3的三个盒子中，用随机变量X放入1号盒的球数, 用随机变量Y表示放入2号盒的球数, 判断X和Y是否独立

二维连续型随机变量

设二维随机变量的分布函数为$F(x,y)$，如果存在二元非负可积函数$f(x,y)$使得对任意实数对(x,y)有

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

则称(X,Y)为二维连续型随机变量，称$\boldsymbol{f(x,y)}$**称为二维随机变量**$\boldsymbol{(X,Y)}$**的概率密度**，或称为**随机变量**$\boldsymbol{X}$**和**$\boldsymbol{Y}$**的联合概率密度**

概率密度函数的性质

- 非负性：$f(x,y) \geq 0$;
- 规范性：$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(u,v)dudv$
- 若$f(x,y)$ 在(x,y) 连续, 则$f(x,y) = \partial^2 F(x,y)/\partial x\partial y$
- 若G为平面上的一个区域, 则点(X,Y)落入G的概率为

$$P((X,Y) \in G) = \iint_{(x,y)\in G} f(x,y)dxdy$$

例题

设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases} ce^{-(3x+4y)} & x>0, y>0 \\ 0 & \text{其它} \end{cases}$$

求$P(0<X<1, 0<Y<2)$

课题练习

设二维随机变量(X,Y)的概率密度

$$f(x,y)=\begin{cases} x^2+axy & 0\le x\le 1,0\le y\le 2 \\ 0 & \text{其它} \end{cases}$$

求$P(X+Y\ge 1)$

边缘概率密度

将随机变量X和Y分别看，依然为随机变量，根据随机变量X的边缘分布函数为

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

$$= \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f(t, y) dt dy = \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} f(t, y) dy \right) dt$$

由此可得随机变量X的边缘概率密度为

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

边缘概率密度

设二维随机变量(X,Y)的概率密度为f(x,y)，则随机变量X和Y的边缘概率密度为

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

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

例题

设二维随机变量(X,Y)的概率密度

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

求$P(X\le 1/2)$