概率论与数理统计

庄玮玮
weizh@ustc.edu.cn

安徽　合肥

2020 年 3 月

第二章　随机变量及概率分布

§2.4 随机向量

在全班的n个人中任选一人，用X，Y分别表示被选人的身高和体重，则X，Y 都是随机变量．这时称(X, Y)是2维随机向量．一般来讲，如果X, Y是定义在同一个概率空间上的随机变量，则称(X, Y) 是2维随机向量，简称为随机向量(random vector).

对随机事件$A, B, A_1, A_2, \cdots, A_n$，以后用$\{A, B\}$ 表示AB，用

$$\{A_1, A_2, \cdots, A_n\} \text{ 表示 } \bigcap_{j=1}^{n} A_j.$$

于是有

$$\{X \leqslant x, Y \leqslant y\} = \{X \leqslant x\} \cap \{Y \leqslant y\}.$$

§2.4 随机向量

在全班的n个人中任选一人，用X，Y分别表示被选人的身高和体重，则X，Y 都是随机变量．这时称(X, Y)是2维随机向量．一般来讲，如果X, Y是定义在同一个概率空间上的随机变量，则称(X, Y) 是2维随机向量，简称为随机向量(random vector).

对随机事件$A, B, A_1, A_2, \cdots, A_n$，以后用$\{A, B\}$ 表示AB，用

$$\{A_1, A_2, \cdots, A_n\} \text{ 表示 } \bigcap_{j=1}^{n} A_j.$$

于是有

$$\{X \leqslant x,\ Y \leqslant y\} = \{X \leqslant x\} \cap \{Y \leqslant y\}.$$

§2.4 随机向量

对于随机变量$X_1, X_2, \cdots, X_m$, 有

$$\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_m \leqslant x_m\} = \bigcap_{j=1}^{m} \{X_j \leqslant x_j\}.$$

对于随机向量(X, Y), 称

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

为(X, Y)的联合分布函数(joint distribution function).

§2.4 随机向量

因为对于$x_1 < x_2$，有

$$\{X \leqslant x_1, Y \leqslant y\} \subset \{X \leqslant x_2, Y \leqslant y\},$$

所以有

$$F(x_1, y) = P(X \leqslant x_1, Y \leqslant y) \leqslant P(X \leqslant x_2, Y \leqslant y) = F(x_2, y).$$

说明联合分布函数$F(x, y)$是x的单调不减函数. 同理，$F(x, y)$也是y的单调不减函数.

§2.4 随机向量

设$F(x, y)$是(X, Y)的联合分布函数，由于$\{Y \leqslant \infty\}$和$\{X \leqslant \infty\}$是必然事件. 所以X, Y分别有概率分布

$$\begin{aligned} F_X(x) &= P(X \leqslant x, Y \leqslant \infty) = F(x, \infty), \\ F_Y(y) &= P(X \leqslant \infty, Y \leqslant y) = F(\infty, y). \end{aligned}$$

这时称X的分布函数$F_X(x)$, Y的分布函数$F_Y(y)$为(X, Y)的边缘分布函数(marginal distribution function).

§2.4 随机向量

定义2.4.1 如果对任何实数x, y, 事件$\{X \leqslant x\}$ 和$\{Y \leqslant y\}$ 独立, 则称随机变量X, Y独立.

X, Y 独立的充分必要条件是对任何 x, y,

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

或等价地有

$$F(x, y) = F_X(x)F_Y(y).$$

下面把随机变量的独立性定义推广到多个随机变量的情况.

§2.4 随机向量

定义2.4.1 如果对任何实数x, y, 事件$\{X \leqslant x\}$ 和$\{Y \leqslant y\}$ 独立, 则称随机变量X, Y独立.

X, Y 独立的充分必要条件是对任何 x, y,

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

或等价地有

$$F(x, y) = F_X(x)F_Y(y).$$

下面把随机变量的独立性定义推广到多个随机变量的情况.

§2.4　随机向量

定义2.4.2　设$X_1, X_2, \cdots$ 是随机变量.

(1) 如果对任何实数$x_1, x_2, \cdots, x_n$,

$$\begin{aligned} & P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n) \\ = \quad & P(X_1 \leqslant x_1)P(X_2 \leqslant x_2)\cdots P(X_n \leqslant x_n), \end{aligned}$$

则称随机变量$X_1, X_2, \cdots, X_n$相互独立.

(2) 如果对任何n, $X_1, X_2, \cdots, X_n$相互独立, 则称随机变量序列$\{X_j\} = \{X_j \mid j = 1, 2, \cdots\}$ 相互独立.

容易理解, 当$X_1, X_2, \cdots, X_n$ 是分别来自相互独立进行的试验的随机变量时, 它们相互独立.

值得指出, 常数与任何随机变量独立. 这是因为任何随机变量的取值都不会影响常数的取值.

§2.4 随机向量

定义2.4.2 设$X_1, X_2, \cdots$ 是随机变量.

(1) 如果对任何实数$x_1, x_2, \cdots, x_n$,

$$\begin{aligned} & P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n) \\ = \quad & P(X_1 \leqslant x_1) P(X_2 \leqslant x_2) \cdots P(X_n \leqslant x_n), \end{aligned}$$

则称随机变量$X_1, X_2, \cdots, X_n$相互独立.

(2) 如果对任何n, $X_1, X_2, \cdots, X_n$相互独立, 则称随机变量序列$\{X_j\} = \{X_j \mid j = 1, 2, \cdots\}$ 相互独立.

容易理解, 当$X_1, X_2, \cdots, X_n$ 是分别来自相互独立进行的试验的随机变量时, 它们相互独立.

值得指出, 常数与任何随机变量独立. 这是因为任何随机变量的取值都不会影响常数的取值.

§2.4 随机向量

在一个城市进行家庭年均收入调查时, 随机选定了n个家庭. 用X_i表示其中第i 个家庭的收入时, X_i是随机变量. 容易理解, 随机变量$X_1, X_2, \cdots, X_n$相互独立. 对于区间 $A_1 = (a_1, b_1]$, $A_2 = (a_2, b_2]$, $\cdots$, $A_n = (a_n, b_n]$, 因为 $X_i \in A_i$ 表示第 i 个家庭的收入在范围 $(a_i, b_i]$ 内, 所以事件

$$\{X_1 \in A_1\},\ \{X_2 \in A_2\},\ \cdots,\ \{X_n \in A_n\}$$

相互独立.

§2.4　随机向量

用Y_i表示第i个家庭在这一年中用于日常生活的支出. 因为支出依赖于收入, 所以Y_i 是X_i 的函数, 可以写成

$$Y_i = g_i(X_i),$$

其中$g_i(x)$是某个函数. 可以理解, 随机变量$Y_1, Y_2, \cdots, Y_n$也相互独立. 若用

$$\overline{X}_k = \frac{X_1 + X_2 + \cdots + X_k}{k}$$

表示前k个家庭的平均收入, 则随机变量$\overline{X}_k$, X_{k+1}, X_{k+2}, $\cdots$, X_n 相互独立.

§2.4 随机向量

定理2.4.1 设 $X_1, X_2, \cdots, X_n$ 相互独立, 则有如下的结果:

- 对于数集 $A_1, A_2, \cdots, A_n$, 事件

$$\{X_1 \in A_1\},\ \{X_2 \in A_2\},\ \cdots,\ \{X_n \in A_n\}$$

  相互独立;

- 对于一元函数 $g_1(x), g_2(x), \cdots, g_n(x)$, 随机变量 $Y_1 = g_1(X_1)$, $Y_2 = g_2(X_2), \cdots, Y_n = g_n(X_n)$ 相互独立;
- 对于 k 元函数 $g(x_1, x_2, \cdots, x_k)$, 定义 $Z_k = g(X_1, X_2, \cdots, X_k)$, 则 $Z_k, X_{k+1}, \cdots, X_n$ 相互独立.

如果 $X_1, X_2, \cdots, X_n$ 都是随机变量, 则称 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 是 n 维随机向量, 也简称为随机向量.

§2.4 随机向量

定义2.4.3 设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 是随机向量，称 R^n 上的 n 元函数

$$F(x_1,x_2,\cdots,x_n)=P(X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n)$$

为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的 联合分布函数. 如果随机向量$\boldsymbol{X}$和随机向量 $\boldsymbol{Y}=(Y_1,Y_2,\cdots,Y_n)$ 有相同的联合分布函数，则称 $\boldsymbol{X}$, $\boldsymbol{Y}$ 同分布.

§2.4　随机向量

设随机向量 $(X_1, X_2, \cdots, X_n)$ 有联合分布函数 $F(x_1, x_2, \cdots, x_n)$, X_i 有分布函数$F_i(x_i)$. 根据独立性的定义知道, $X_1, X_2, \cdots, X_n$ 相互独立的充分必要条件是对任何 $(x_1, x_2, \cdots, x_n)$, 有

$$F(x_1, x_2, \cdots, x_n) = F_1(x_1)F_2(x_2)\cdots F_n(x_n).$$

容易理解, 如果 S_1, S_2, $\cdots$, S_n 是 n 个独立进行的试验, X_i 是试验 S_i 下的随机变量, 则 X_1, X_2, $\cdots$, X_n 相互独立. 如果 S_1, S_2, $\cdots$ 是独立进行的试验, X_i 是试验 S_i 下的事件, 则 X_1, X_2, $\cdots$ 相互独立.

§2.4 随机向量

§2.4.1 离散型随机向量

如果X, Y都是离散型随机变量, 则称(X, Y)是离散型随机向量. 设离散型随机向量(X, Y) 有联合概率分布

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j \geqslant 1, \qquad (2.4.1)$$

则X和Y分别有概率分布

$$\begin{aligned} p_i \equiv P(X = x_i) &= \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{j=1}^{\infty} p_{ij}, \ i \geqslant 1, \\ q_j \equiv P(Y = y_j) &= \sum_{i=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{i=1}^{\infty} p_{ij}, \ j \geqslant 1. \end{aligned}$$

这时称X的分布$\{p_i\}$, Y的分布$\{q_j\}$为(X, Y)的边缘分布.

§2.4 随机向量

§2.4.1 离散型随机向量

联合概率分布常被简称为联合分布或概率分布.

当(X,Y)的联合分布的规律性不强，或不能用(2.4.1)明确表达时，还可以用表格的形式表达如下:

p_{ij}	y_1	y_2	y_3	$\cdots$	y_n	$\cdots$	$\{p_i\}$
x_1	p_{11}	p_{12}	p_{13}	$\cdots$	p_{1n}	$\cdots$	p_1
x_2	p_{21}	p_{22}	p_{23}	$\cdots$	p_{2n}	$\cdots$	p_2
x_3	p_{31}	p_{32}	p_{33}	$\cdots$	p_{3n}	$\cdots$	p_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
$\{q_j\}$	q_1	q_2	q_3	$\cdots$	q_n	$\cdots$	1

其中$p_i = P(X = x_i)$是其所在行中的诸p_{ij}之和，$q_j = P(Y = y_j)$是其所在列的诸p_{ij} 之和.

§2.4 随机向量

§2.4.1 离散型随机向量

回忆定义2.4.1: 如果对任何实数x, y, 有

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

则称随机变量X, Y独立.

关于离散型的(X, Y), 我们有如下的定理.

定理2.4.1 设离散型随机向量(X, Y)的所有不同取值是

$$(x_i, y_j), \quad i, j \geqslant 1,$$

则X, Y独立的充分必要条件是对任何(x_i, y_j),

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j). \qquad (2.4.2)$$

§2.4　随机向量

§2.4.1　离散型随机向量

回忆定义2.4.1: 如果对任何实数x, y, 有

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

则称随机变量X, Y独立.

关于离散型的(X, Y), 我们有如下的定理.

定理2.4.1　设离散型随机向量(X, Y)的所有不同取值是

$$(x_i, y_j), \quad i, j \geqslant 1,$$

则X, Y独立的充分必要条件是对任何(x_i, y_j),

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j). \tag{2.4.2}$$

§2.4　随机向量

§2.4.1　离散型随机向量

解　当(2.4.2)成立，对任何x, y，在(2.4.2)的两边对$\{i\,|x_i \leqslant x\}$中的i求和，得到

$$P(X \leqslant x, Y = y_j) = P(X \leqslant x)P(Y = y_j).$$

再对 $\{j\,|y_j \leqslant y\}$ 中的 j 求和，得到

$$P(X \leqslant x, Y \leqslant y) = P(X \leqslant x)P(Y \leqslant y).$$

于是 X，Y 独立.

反之，设 X，Y 独立. 对于单点集合 $A = \{x_i\}$ 和 $B = \{y_j\}$，由定理2.4.1知道 $\{X = x_i\} = \{X \in A\}$ 与 $\{Y = y_j\} = \{Y \in B\}$ 独立，所以(2.4.2)成立.

§2.4 随机向量

§2.4.1 离散型随机向量

例2.4.1(接§2.1 例2.1.2) 设一部手机在时间段$[0, t]$内收到的短信数服从泊松分布$\mathcal{P}(\lambda)$, 其中$\lambda = \mu t$, μ是正常数. 每个短信是否广告与其到达时间独立, 也与其他短信是否广告独立. 如果每个短信是广告的概率为正数p, 则$[0, t]$内到达的广告短信数和非广告短信数相互独立.

§2.4 随机向量

§2.4.1 离散型随机向量

证明 设 $[0,t]$ 内收到的短信数是 Y. 根据§2.1例2.1.2 的结论, 收到的广告短信数 $X \sim \mathcal{P}(\lambda p)$, 收到的非广告短信数 $Z = Y - X \sim \mathcal{P}(\lambda q)$, 其中 $q = 1 - p$. 每收到一个短信相当于作一次试验, 遇到广告是试验成功. 于是得到

$$
\begin{aligned}
P(X=k, Z=j) &= P(X=k, Y-X=j) \\
&= P(X=k, Y=j+k) \\
&= P(Y=j+k)P(X=k|Y=j+k) \\
&= \frac{\lambda^{j+k}}{(j+k)!}\mathrm{e}^{-\lambda}\mathrm{C}_{j+k}^{k}p^k q^j \\
&= \frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda p}\,\frac{(\lambda q)^j}{j!}\mathrm{e}^{-\lambda q} \\
&= P(X=k)P(Z=j).
\end{aligned}
$$

从定理2.4.1知道 X, Y 独立.

§2.4　随机向量

§2.4.1　离散型随机向量

例2.4.2(三项分布)　设A, B, C是试验S的完备事件组，$P(A) = p_1$，$P(B) = p_2$，$P(C) = p_3$. 对试验S进行n次独立重复试验时，用X_1, X_2, X_3分别表示A, B, C发生的次数，则(X_1, X_2, X_3) 的联合分布是

$$P(X_1 = i, X_2 = j, X_3 = k) = \frac{n!}{i!\, j!\, k!} p_1^i p_2^j p_3^k,$$

其中$i, j, k \geqslant 0$，$i + j + k = n$.

§2.4 随机向量

§2.4.1 离散型随机向量

证明 将$1,2,\cdots,n$分成有次序的3组，不考虑每组中元素的次序，第 $1,2,3$ 组分别有 i,j,k 个元素的不同结果共有

$$N=\frac{n!}{i!\,j!\,k!}$$

个．用第l个分组结果 $\{a_1,a_2,\cdots,a_i\}\equiv A_l$，$\{b_1,b_2,\cdots,b_j\}\equiv B_l$，$\{c_1,c_2,\cdots,c_k\}\equiv C_l$ 表示第 $a_1,a_2,\cdots,a_i$ 次试验 A 发生，第$b_1,b_2,\cdots,b_j$次试验B发生，第$c_1,c_2,\cdots,c_k$次试验C发生，则

$$P(A_lB_lC_l)=p_1^ip_2^jp_3^k,\ 1\leqslant l\leqslant N.$$

因为对于不同的l，事件$A_lB_lC_l$ 互不相容，所以得到

$$\begin{aligned}P(X_1=i,X_2=j,X_3=k) &= P\Big(\bigcup_{l=1}^{N}A_lB_lC_l\Big)\\ &= \sum_{l=1}^{N}P\big(A_lB_lC_l\big)=\frac{n!}{i!\,j!\,k!}p_1^ip_2^jp_3^k.\end{aligned}$$

§2.4 随机向量

§2.4.1 离散型随机向量

例2.4.3 在例2.4.2中，计算 X_1 和 (X_1, X_2) 的边缘分布.

解 在例2.4.2中的单次试验中，如果A发生就称试验成功，则试验成功的概率是$p_1 = P(A)$. X_1是n次独立重复试验中成功的次数，所以X_1有概率分布

$$P(X_1 = i) = \mathrm{C}_n^i p_1^i (1 - p_1)^{n-i}, \quad i \geqslant 0.$$

即 $X_1 \sim \mathcal{B}(n, p_1)$. 因为 $X_1 + X_2 + X_3 = n$, 所以

$$\begin{aligned} P(X_1 = i, X_2 = j) &= P(X_1 = i, X_2 = j, X_3 = n - i - j) \\ &= \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j}, \quad i + j \leqslant n. \end{aligned}$$

例2.4.3告诉我们，在一些情况下根据问题的背景更容易计算出随机向量的边缘分布.

§2.4 随机向量

§2.4.1 离散型随机向量

例2.4.3 在例2.4.2中，计算 X_1 和 (X_1, X_2) 的边缘分布.

解 在例2.4.2中的单次试验中，如果A发生就称试验成功，则试验成功的概率是$p_1 = P(A)$. X_1是n次独立重复试验中成功的次数，所以X_1有概率分布

$$P(X_1 = i) = \mathrm{C}_n^i p_1^i (1 - p_1)^{n-i}, \quad i \geqslant 0.$$

即 $X_1 \sim \mathcal{B}(n, p_1)$. 因为 $X_1 + X_2 + X_3 = n$, 所以

$$\begin{aligned} P(X_1 = i, X_2 = j) &= P(X_1 = i, X_2 = j, X_3 = n - i - j) \\ &= \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j}, \quad i + j \leqslant n. \end{aligned}$$

例2.4.3告诉我们，在一些情况下根据问题的背景更容易计算出随机向量的边缘分布.

§2.4 随机向量

§2.4.1 离散型随机向量

例2.4.4(多项分布) 设$A_1, A_2, \cdots, A_r$是试验S的完备事件组. 对试验 S 进行 n 次独立重复试验时, 用 X_i 表示 A_i 发生的次数, 则 $(X_1, X_2, \cdots, X_r)$ 的联合分布是

$$\begin{aligned} & P(X_1 = k_1, X_2 = k_2, \cdots, X_r = k_r) \\ = & \frac{n!}{k_1!k_2!\cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}, \end{aligned}$$

其中 $k_i \geqslant 0$, $p_i = P(A_i)$, $\sum_{i=1}^{r} k_i = n$.

§2.4 随机向量

§2.4.2 连续型随机向量

A. 联合密度

在向平面坐标系的原点射击时，用 (X, Y) 表示弹落点，(X, Y) 是随机向量. 对于任何指定的常数向量 (x, y) 都有

$$P(X=x, Y=y) \leqslant P(X=x)=0.$$

但是对于任何包含原点且面积为正数的长方形

$$D=\{ (x, y) \mid a<x \leqslant b, c<y \leqslant d \},$$

有

$$P((X, Y) \in D)>0.$$

为了刻画上述概率，可以设想有一个 (x, y) 平面上的曲面 $z=f(x, y)$，使得概率 $P((X, Y) \in D)$ 等于以 D 为底，$f(x, y)$ 为顶的柱体体积. 于是有下面的定义.

§2.4 随机向量

§2.4.2 连续型随机向量

定义2.4.3 设 (X, Y) 是随机向量，如果有 R^2 上的非负函数 $f(x, y)$ 使得对 R^2 的任何长方形子集

$$D = \{\ (x, y) \mid a < x \leqslant b, c < y \leqslant d\ \}$$

有

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

则称 (X, Y) 是连续型随机向量，并称 $f(x, y)$ 是 (X, Y) 的联合概率密度或联合密度(joint density).

按照上述定义，连续型随机向量有联合密度，没有联合密度的随机向量不是连续型随机向量. 另外，如果两个随机向量有相同的联合密度，则称它们同分布.

§2.4 随机向量

§2.4.2 连续型随机向量

定义2.4.3 设 (X, Y) 是随机向量，如果有 R^2 上的非负函数 $f(x, y)$ 使得对 R^2 的任何长方形子集

$$D = \{ (x, y) \mid a < x \leqslant b, c < y \leqslant d \}$$

有

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

则称 (X, Y) 是连续型随机向量，并称 $f(x, y)$ 是 (X, Y) 的联合概率密度或联合密度(joint density).

按照上述定义，连续型随机向量有联合密度，没有联合密度的随机向量不是连续型随机向量. 另外，如果两个随机向量有相同的联合密度，则称它们同分布.

§2.4 随机向量

§2.4.2 连续型随机向量

设 $f(x,y)$ 是 (X,Y) 的联合密度. 可以证明对 R^2!的任何子区域!B, 有

$$
\begin{aligned}
P((X,Y)\in B) &= \iint_{R^2} \mathrm{I}[(x,y)\in B] f(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \iint_{B} f(x,y)\,\mathrm{d}x\mathrm{d}y,
\end{aligned}
$$

其中

$$
\mathrm{I}[(x,y)\in B] = \begin{cases} 1, & \text{当}\,(x,y)\in B, \\ 0, & \text{否则} \end{cases} \qquad (2.4.3)
$$

是集合B的示性函数, 也常简写成$\mathrm{I}[B]$.

§2.4 随机向量

§2.4.2 连续型随机向量

于是

$$\mathrm{I}[B] = \mathrm{I}[(x, y) \in B].$$

公式(2.4.3)是常用公式, 值得牢记. 在(2.4.3)中取 $B = R^2$ 时, 得到

$$\iint_{R^2} f(x, y)\,\mathrm{d}x\mathrm{d}y = P((X, Y) \in R^2) = 1.$$

为了计算重积分的方便, 列出下面的定理.

§2.4　随机向量

§2.4.2　连续型随机向量

定理2.4.2　设D是R^2的子区域，函数$h(x,y)$在D中非负，或$|h(x,y)|$在D'上的积分有限. 用$\mathrm{I}[D]$表示D的示性函数，则

$$\begin{aligned}\iint_D h(x,y)\,\mathrm{d}x\mathrm{d}y &= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} h(x,y)\mathrm{I}[D]\,\mathrm{d}y\Big)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} h(x,y)\mathrm{I}[D]\,\mathrm{d}x\Big)\,\mathrm{d}y.\end{aligned}$$

§2.4 随机向量

§2.4.2 连续型随机向量

定理2.4.2 给出了化二重积分为一元积分的方法. 注意, 计算

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} h(x,y)\mathrm{I}[D]\,\mathrm{d}y\right)\mathrm{d}x$$

时, 将x视为常数先对y积出$\int_{-\infty}^{\infty} h(x,y)\mathrm{I}[D]\,\mathrm{d}y$, 然后再对$x$ 进行积分. 同理, 计算

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} h(x,y)\mathrm{I}[D]\,\mathrm{d}x\right)\mathrm{d}y$$

时, 将y视为常数先对x积出$\int_{-\infty}^{\infty} h(x,y)\mathrm{I}[D]\,\mathrm{d}x$, 然后再对$y$ 进行积分.

§2.4 随机向量

§2.4.2 连续型随机向量

B. 边缘密度

如果 $f(x,y)$ 是随机向量 (X,Y) 的联合密度，则称 X, Y 各自的概率密度为 $f(x,y)$ 或 (X,Y) 的边缘密度 (marginal density)，下面计算 (X,Y) 的边缘密度. 对任何x，从概率密度的定义和

$$\begin{aligned} P(X \leqslant x) &= P(X \leqslant x, Y < \infty) \\ &= \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \end{aligned}$$

知道X有边缘密度

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y. \qquad (2.4.4)$$

§2.4　随机向量

§2.4.2　连续型随机向量

完全对称地得到Y的边缘函数

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

从联合密度计算边缘密度的公式是容易掌握的: 求$f_X(x)$ 时, 对$f(x,y)$的y积分, 留下x; 求$f_Y(y)$时, 对$f(x,y)$的x积分, 留下y.
设 D 是 R^2 的子区域, D 的面积 $m(D)$ 是正数. 如果 (X, Y) 有联合密度

$$f(x, y) = \begin{cases} \frac{1}{m(D)}, & (x, y) \in D, \\ 0, & (x, y) \notin D, \end{cases}$$

则称 (X, Y) 在 D 上均匀分布, 记做 $(X, Y) \sim \mathcal{U}(D)$.

§2.4　随机向量

§2.4.2　连续型随机向量

例2.4.5　设 (X, Y) 在单位圆 $D = \{(x, y)|x^2 + y^2 \leqslant 1\}$ 内均匀分布，求 X 和 Y 的概率密度.

解　用 $\mathrm{I}[D]$ 表示 D 的示性函数，即

$$\mathrm{I}[D] = \mathrm{I}[x^2 + y^2 \leqslant 1] = \begin{cases} 1, & \text{当} x^2 + y^2 \leqslant 1, \\ 0, & \text{否则}, \end{cases}$$

则(X, Y)有联合密度 $f(x, y) = (1/\pi)\mathrm{I}[D]$. X 只在 $[-1, 1]$ 中取值.

§2.4 随机向量

§2.4.2 连续型随机向量

例2.4.5 设 (X, Y) 在单位圆 $D = \{(x, y) | x^2 + y^2 \leqslant 1\}$ 内均匀分布, 求 X 和 Y 的概率密度.

解 用 $\mathrm{I}[D]$ 表示 D 的示性函数, 即

$$\mathrm{I}[D] = \mathrm{I}[x^2 + y^2 \leqslant 1] = \begin{cases} 1, & \text{当} x^2 + y^2 \leqslant 1, \\ 0, & \text{否则}, \end{cases}$$

则(X, Y)有联合密度 $f(x, y) = (1/\pi)\mathrm{I}[D]$. X 只在 $[-1, 1]$ 中取值.

§2.4 随机向量

§2.4.2 连续型随机向量

由(2.4.4)知道

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{I}[x^2+y^2 \leqslant 1]\,\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{I}[\,|y| \leqslant \sqrt{1-x^2}\,]\,\mathrm{d}y \\
&= \frac{2}{\pi}\sqrt{1-x^2}, \quad |x| \leqslant 1.
\end{aligned}
$$

对称地得到 Y 的概率密度

$$
f_Y(y) = (2/\pi)\sqrt{1-y^2}, \ |y| \leqslant 1.
$$

§2.4 随机向量

§2.4.2 连续型随机向量

C. 独立性

关于连续型随机变量的独立性, 我们介绍下面的定理.

定理2.4.3 设 X, Y 分别有概率密度 $f_X(x)$, $f_Y(y)$. 则 X, Y 独立的充分必要条件是随机向量 (X, Y) 有联合密度

$$f(x,y) = f_X(x)f_Y(y). \qquad (2.4.5)$$

证明 如果(2.4.5)是 (X, Y) 的联合密度, 则有

$$\begin{aligned} P(X \leqslant x, Y \leqslant y) &= \int_{-\infty}^{x} \left(\int_{-\infty}^{y} f_X(s) f_Y(t) \, \mathrm{d}t \right) \mathrm{d}s \\ &= \int_{-\infty}^{x} f_X(s) \, \mathrm{d}s \int_{-\infty}^{y} f_Y(t) \, \mathrm{d}t \\ &= P(X \leqslant x) P(Y \leqslant y). \end{aligned}$$

由定义, 可知 X, Y 独立.

§2.4 随机向量

§2.4.2 连续型随机向量

如果 X, Y 独立, 对 $a \leqslant b,\ c \leqslant d$, 利用定理2.4.2 得到

$$\begin{aligned}
& P(a < X \leqslant b, c < Y \leqslant d) \\
= & P(a < X \leqslant b)P(c < Y \leqslant d) \\
= & \int_a^b f_X(x)\,\mathrm{d}x \int_c^d f_Y(y)\,\mathrm{d}y \\
= & \int_a^b \int_c^d f_X(x) f_Y(y)\,\mathrm{d}x\mathrm{d}y.
\end{aligned}$$

从联合密度的定义知道 $f_X(x)f_Y(y)$ 是 (X, Y) 的联合密度.

§2.4 随机向量

§2.4.2 连续型随机向量

设 X 有概率密度 $f_X(x)$, 则 X 的取值范围是 $\{x|f_X(x)>0\}$. 如果观测到 $X=x$, 则 $f_X(x)>0$. 设(X,Y)有联合密度$f(x,y)$, 对于确定的x, 已知$X=x$时, Y的取值范围是

$$\{y\,|f(x,y)>0\}.$$

如果 X,Y 独立, 则已知 $X=x$ 时, 可将 Y 的取值范围写成

$$\{y\,|f_X(x)f_Y(y)>0\}=\{y\,|f_Y(y)>0\}.$$

上式右边与 x 无关. 说明如果 X,Y 独立, 则已知 $X=x$ 时, Y 的取值范围与 x 无关.

§2.4 随机向量

§2.4.2 连续型随机向量

定理2.4.4 设 (X, Y) 是随机向量. 已知 $X = x$ 时, 如果 Y 的取值范围和 x 有关, 则 X, Y 不独立.

例2.4.6 设(X, Y)在矩形 $D = \{(x, y) \mid a < x \leqslant b, c < y \leqslant d\}$ 上均匀分布. 计算 X, Y 的边缘分布, 并证明 X, Y 独立.

解 用 $\mathrm{I}[D]$ 表示 D 的示性函数, 则

$$\mathrm{I}[D] = \mathrm{I}[a < x \leqslant b] \cdot \mathrm{I}[c < y \leqslant d],$$

(X, Y) 的联合密度

$$\frac{1}{m(D)}\mathrm{I}[D] = \frac{1}{b-a}\mathrm{I}[a < x \leqslant b] \cdot \frac{1}{d-c}\mathrm{I}[c < y \leqslant d].$$

§2.4 随机向量

§2.4.2 连续型随机向量

定理2.4.4 设 (X,Y) 是随机向量. 已知 $X=x$ 时, 如果 Y 的取值范围和 x 有关, 则 X,Y 不独立.

例2.4.6 设(X,Y)在矩形 $D=\{(x,y)\,|a<x\leqslant b, c<y\leqslant d\}$ 上均匀分布. 计算 X, Y 的边缘分布, 并证明 X, Y 独立.

解 用 $\mathrm{I}[D]$ 表示 D 的示性函数, 则

$$\mathrm{I}[D]=\mathrm{I}[a<x\leqslant b]\cdot\mathrm{I}[c<y\leqslant d],$$

(X,Y) 的联合密度

$$\frac{1}{m(D)}\mathrm{I}[D]=\frac{1}{b-a}\mathrm{I}[a<x\leqslant b]\cdot\frac{1}{d-c}\mathrm{I}[c<y\leqslant d].$$

§2.4 随机向量

§2.4.2 连续型随机向量

容易计算出 X 和 Y 的概率密度如下:

$$f_X(x) = \frac{1}{b-a}\mathrm{I}[a < x \leqslant b], \quad f_Y(y) = \frac{1}{d-c}\mathrm{I}[c < y \leqslant d].$$

于是 $X \sim \mathcal{U}(a,b]$, $Y \sim \mathcal{U}(c,d]$. 由于 $f_X(x)f_Y(y) = f(x,y)$, 所以 X, Y 相互独立.

现在将例2.4.6 中的矩形 D 作一转动, 使得矩形的边不与坐标轴平行. 这时 X 的取值会影响 Y 的取值范围, 因而 X, Y 不再独立. 同理, 如果 (X,Y) 的联合密度仅在圆、椭圆或三角形内大于0, 则 X, Y 不独立.

§2.4 随机向量

§2.4.2 连续型随机向量

例2.4.7 两人某天在1点至2点间独立地随机到达某地会面，先到者等候20 min 后离去．求这两人能相遇的概率．

解 认为每个人在0至60 min内等可能到达，用X，Y 分别表示他们的到达时间．则$X \sim \mathcal{U}(0,60)$, $Y \sim \mathcal{U}(0,60)$, X, Y独立．利用

$$f_X(x) = \begin{cases} \frac{1}{60}, & x \in (0,60), \\ 0, & x \notin (0,60), \end{cases} \qquad f_Y(y) = \begin{cases} \frac{1}{60}, & y \in (0,60), \\ 0, & y \notin (0,60), \end{cases}$$

得到 (X, Y) 的联合密度

$$f(x,y) = f_X(x) f_Y(y) = \begin{cases} 1/60^2, & (x,y) \in D, \\ 0, & (x,y) \notin D, \end{cases}$$

其中$D = \{(x,y) | 0 \leqslant x, y \leqslant 60\}$.

§2.4 随机向量

§2.4.2 连续型随机向量

例2.4.7 两人某天在1点至2点间独立地随机到达某地会面，先到者等候20 min 后离去. 求这两人能相遇的概率.

解 认为每个人在0至60 min内等可能到达，用X，Y 分别表示他们的到达时间. 则$X \sim \mathcal{U}(0,60)$，$Y \sim \mathcal{U}(0,60)$，X, Y独立. 利用

$$f_X(x) = \begin{cases} \frac{1}{60}, & x \in (0,60), \\ 0, & x \notin (0,60), \end{cases} \quad f_Y(y) = \begin{cases} \frac{1}{60}, & y \in (0,60), \\ 0, & y \notin (0,60), \end{cases}$$

得到 (X, Y) 的联合密度

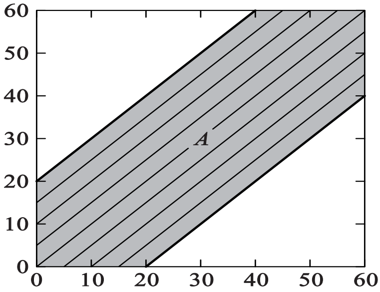
$$f(x,y) = f_X(x) f_Y(y) = \begin{cases} 1/60^2, & (x,y) \in D, \\ 0, & (x,y) \notin D, \end{cases}$$

其中$D = \{(x,y) | 0 \leqslant x, y \leqslant 60\}$.

§2.4 随机向量

§2.4.2 连续型随机向量

$$A = \{\ (x,y)\ \big|\ |x-y| \leqslant 20, (x,y) \in D\ \}.$$



§2.4　随机向量

§2.4.2　连续型随机向量

要计算的概率是

$$
\begin{aligned}
P(|X-Y|\leqslant 20) &= \iint_A f(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \frac{m(A)}{m(D)} \\
&= \frac{60^2-40^2}{60^2} \\
&= \frac{5}{9}.
\end{aligned}
$$

§2.4 随机向量

§2.4.2 连续型随机向量

例2.4.8 设(X, Y)在由曲线 $y = x^2/2$ 和 $y = x$ 所围的有限区域内均匀分布.

(a) 求 (X, Y) 的联合密度;

(b) X, Y 是否独立;

(c) 计算 $f_X(x)$ 和 $f_Y(y)$.

解 从$x^2/2 = x$ 解出 $x = 0$ 或 2. 两条曲线的交点是 $(0, 0)$ 和 $(2, 2)$, 所述的区域是:

$$D = \{(x, y) \,|\, x^2/2 < y < x, 0 < x < 2\}.$$

§2.4 随机向量

§2.4.2 连续型随机向量

例2.4.8 设(X, Y)在由曲线 $y = x^2/2$ 和 $y = x$ 所围的有限区域内均匀分布.

(a) 求 (X, Y) 的联合密度;

(b) X, Y 是否独立;

(c) 计算 $f_X(x)$ 和 $f_Y(y)$.

解 从$x^2/2 = x$ 解出 $x = 0$ 或 2. 两条曲线的交点是 $(0, 0)$ 和 $(2, 2)$, 所述的区域是:

$$D = \{(x, y) \mid x^2/2 < y < x, 0 < x < 2\}.$$

§2.4　随机向量

§2.4.2　连续型随机向量

用$\mathrm{I}[D]=\mathrm{I}[x^2/2<y<x]\cdot\mathrm{I}[0<x<2]$表示$D$的示性函数，则$D$的面积是

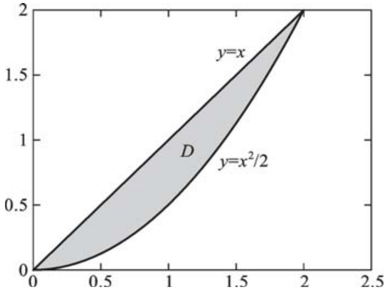
$$
\begin{aligned}
m(D) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{I}[D]\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty}\mathrm{I}[x^2/2<y<x]\,\mathrm{d}y\Big)\mathrm{I}[0<x<2]\,\mathrm{d}x \\
&= \int_{0}^{2}\Big(\int_{x^2/2}^{x}\mathrm{d}y\Big)\mathrm{d}x \\
&= \int_{0}^{2}(x-x^2/2)\,\mathrm{d}x \\
&= \frac{4}{2}-\frac{4}{3}=\frac{2}{3}.
\end{aligned}
$$

§2.4 随机向量

§2.4.2 连续型随机向量

(a) (X, Y)的联合密度是

$$f(x,y)=\frac{1}{m(D)}\mathrm{I}[D]=1.5\,\mathrm{I}[D].$$



§2.4 随机向量

§2.4.2 连续型随机向量

(b) 因为已知$X=x$时，Y在$(x^2/2,x)$中取值. X的取值影响了Y的取值范围，所以X 和Y 不独立.

(c) 由 $\{x^2/2<y<x\}=\{y<x<\sqrt{2y}\}$ 计算出

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} 1.5\,\mathrm{I}[D]\,\mathrm{d}y = 1.5\int_{x^2/2}^{x} \mathrm{I}[0<x<2]\,\mathrm{d}y \\
&= 1.5(x-x^2/2)\mathrm{I}[0<x<2] \\
&= 1.5(x-x^2/2),\ \ x\in(0,2), \\
f_Y(y) &= \int_{-\infty}^{\infty} 1.5\mathrm{I}[D]\,\mathrm{d}x = 1.5\int_{0}^{2} \mathrm{I}[y<x<\sqrt{2y}]\,\mathrm{d}x \\
&= 1.5(\sqrt{2y}-y),\ \ y\in(0,2).
\end{aligned}
$$

§2.4 随机向量

§2.4.2 连续型随机向量

定理2.4.5 若连续型随机向量 $(X_1,\dots,X_n)$ 的概率密度函数 $f(x_1,\dots,x_n)$ 可表为 n 个函数 $g_1,\dots,g_n$ 之积，其中 g_i 只依赖于 x_i，即

$$f(x_1,\dots,x_n)=g_1(x_1)\cdots g_n(x_n),$$

则 $X_1,\dots,X_n$ 相互独立，且 X_i 的边缘密度函数 $f_i(x_i)$ 与 $g_i(x_i)$ 只相差一个常数因子。

例2.4.9 设 X_1, X_2 独立，都服从标准正态分布 $N(0,1)$. 把点 (X_1,X_2) 的极坐标记为 (R,Θ) $(0\le R<\infty, 0\le\Theta<2\pi)$. 求证： R 和 Θ 独立。

§2.4 随机向量

§2.4.2 连续型随机向量

定理2.4.5 若连续型随机向量 $(X_1, \ldots, X_n)$ 的概率密度函数 $f(x_1, \ldots, x_n)$ 可表为 n 个函数 $g_1, \ldots, g_n$ 之积，其中 g_i 只依赖于 x_i，即

$$f(x_1, \ldots, x_n) = g_1(x_1) \cdots g_n(x_n),$$

则 $X_1, \ldots, X_n$ 相互独立，且 X_i 的边缘密度函数 $f_i(x_i)$ 与 $g_i(x_i)$ 只相差一个常数因子。

例2.4.9 设 X_1, X_2 独立，都服从标准正态分布 $N(0, 1)$. 把点 (X_1, X_2) 的极坐标记为 (R, Θ) $(0 \le R < \infty, 0 \le \Theta < 2\pi)$. 求证： R 和 Θ 独立。

§2.5　随机向量函数的分布

先将§2.4中的全概率公式作一推广.

定理2.5.1(全概率公式)　如果随机变量 X 有概率分布 $p_j = P(X = x_j), j \geqslant 0$, 则对事件 B 有

$$P(B) = \sum_{j=0}^{\infty} P(B|X = x_j)P(X = x_j).$$

解　事件 $A_j = \{X = x_j\}, j \geqslant 0$ 构成完备事件组. 所以有

$$\bigcup_{j=0}^{\infty} A_j = \Omega,\ \ B = B\bigcup_{j=0}^{\infty} A_j = \bigcup_{j=0}^{\infty} BA_j.$$

§2.5 随机向量函数的分布

先将§2.4中的全概率公式作一推广.

定理2.5.1(全概率公式) 如果随机变量 X 有概率分布 $p_j = P(X = x_j), j \geqslant 0$, 则对事件 B 有

$$P(B) = \sum_{j=0}^{\infty} P(B|X = x_j)P(X = x_j).$$

解 事件 $A_j = \{X = x_j\}, j \geqslant 0$ 构成完备事件组. 所以有

$$\bigcup_{j=0}^{\infty} A_j = \Omega,\ B = B\bigcup_{j=0}^{\infty} A_j = \bigcup_{j=0}^{\infty} BA_j.$$

§2.5 随机向量函数的分布

因为 $BA_1, BA_2, \cdots$ 互不相容，所以用概率的可加性得到

$$
\begin{aligned}
P(B) &= P\left(\bigcup_{j=0}^{\infty} BA_j\right) \\
&= \sum_{j=0}^{\infty} P(BA_j) \\
&= \sum_{j=0}^{\infty} P(B|A_j)P(A_j).
\end{aligned}
$$

§2.5 随机向量函数的分布

A. 离散型随机向量的函数

下面通过例子学习计算离散型随机向量函数的概率分布.

例2.5.1(泊松分布的可加性) 设一个公交车站有1路, 2路, $\cdots$, n 路汽车停靠. 早7 点至8点之间乘i路车的乘客的到达数X_i服从参数是λ_i的泊松分布. 设$X_1, X_2, \cdots, X_n$相互独立, 计算7 点至8 点之间

(a) 乘1路和2路汽车的到达人数$Z_2 = X_1 + X_2$的概率分布;

(b) 到达总人数$Z_n = X_1 + X_2 + \cdots + X_n$的概率分布.

解 (a) Z_2是取非负整数值的随机变量. 用$B|A$表示已知A发生后的事件B, 则对$k = 0, 1, \cdots$, 有

$$\{Z_2 = k|X_1 = i\} = \{i + X_2 = k|X_1 = i\}.$$

§2.5 随机向量函数的分布

A. 离散型随机向量的函数

下面通过例子学习计算离散型随机向量函数的概率分布.

例2.5.1(泊松分布的可加性) 设一个公交车站有1路, 2路, $\cdots$, n 路汽车停靠. 早7 点至8点之间乘i路车的乘客的到达数X_i服从参数是λ_i的泊松分布. 设$X_1, X_2, \cdots, X_n$相互独立, 计算7 点至8 点之间

(a) 乘1路和2路汽车的到达人数$Z_2 = X_1 + X_2$的概率分布;

(b) 到达总人数$Z_n = X_1 + X_2 + \cdots + X_n$的概率分布.

解 (a) Z_2是取非负整数值的随机变量. 用$B|A$表示已知A发生后的事件B, 则对$k = 0, 1, \cdots$, 有

$$\{Z_2 = k | X_1 = i\} = \{i + X_2 = k | X_1 = i\}.$$

§2.5 随机向量函数的分布

用全概率公式得到

$$
\begin{aligned}
P(Z_2=k) &= \sum_{i=0}^{\infty} P(Z_2=k|X_1=i)P(X_1=i) \\
&= \sum_{i=0}^{k} P(X_2=k-i)P(X_1=i) \\
&= \sum_{i=0}^{k} \frac{\lambda_2^{k-i}}{(k-i)!}\frac{\lambda_1^i}{i!}\mathrm{e}^{-\lambda_1-\lambda_2} \\
&= \frac{1}{k!}\sum_{i=0}^{k} \mathrm{C}_k^i \lambda_1^i \lambda_2^{k-i}\mathrm{e}^{-\lambda_1-\lambda_2} \\
&= \frac{(\lambda_1+\lambda_2)^k}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2)}.
\end{aligned}
$$

说明到达人数$Z_2=X_1+X_2\sim\mathcal{P}(\lambda_1+\lambda_2)$.

§2.5 随机向量函数的分布

(b) 用归纳法. 假设$Z_{n-1} \sim \mathcal{P}(\lambda_1 + \lambda_2 + \cdots + \lambda_{n-1})$. 利用$Z_{n-1}$和$X_n$独立, $Z_n = Z_{n-1} + X_n$和(1)中的结果得到

$$Z_n \sim \mathcal{P}(\lambda_1 + \lambda_2 + \cdots + \lambda_n).$$

从例2.5.1可以得到如下的结果: 如果$X_1, X_2, \cdots, X_n$相互独立, $X_i \sim \mathcal{P}(\lambda_i)$, 则$Z_n = X_1 + X_2 + \cdots + X_n \sim \mathcal{P}(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$.

§2.5 随机向量函数的分布

从问题的背景也可以理解两个相互独立的服从泊松分布的随机变量之和仍然服从泊松分布.

▶ 在放射物放射α粒子的例子中，放射物在0.75秒内放射出的α粒子数服从$\lambda = 3.87$ 的泊松分布. λ是7.5秒内平均放射出的粒子数.

▶设想将此放射物分成两块，则各块放射的粒子数相互独立，都服从泊松分布. 若第1块在7.5秒内平均释放出 λ_1个α 粒子，第2块在7.5秒内平均释放出 λ_2 个 α 粒子，则两块之和在7.5秒内平均释放出
$\lambda = \lambda_1 + \lambda_2$ 个 α 粒子.

§2.5 随机向量函数的分布

例2.5.2 实验室有n个学生, 在相同的条件下每人独立重复同一试验. 如果第i个人作了m_i次试验, 其中试验成功的次数是X_i. 计算这n个学生试验的成功总次数

$$Z_n = X_1 + X_2 + \cdots + X_n$$

的概率分布.

解 设每次试验成功的概率是p. 因为这n个同学一共进行了$m = m_1 + m_2 + \cdots + m_n$ 次独立重复试验, 所以试验成功的总次数 Z_n 服从二项分布$\mathcal{B}(m, p)$.

§2.5 随机向量函数的分布

例2.5.2 实验室有n个学生，在相同的条件下每人独立重复同一试验．如果第i个人作了m_i次试验，其中试验成功的次数是X_i．计算这n个学生试验的成功总次数

$$Z_n = X_1 + X_2 + \cdots + X_n$$

的概率分布．

解 设每次试验成功的概率是p．因为这n个同学一共进行了$m = m_1 + m_2 + \cdots + m_n$ 次独立重复试验，所以试验成功的总次数 Z_n 服从二项分布$\mathcal{B}(m, p)$．

§2.5 随机向量函数的分布

例2.5.2说明: 如果X_i 服从二项分布$\mathcal{B}(m_i,p)$, $X_1,X_2,\cdots,X_n$ 相互独立, 则它们的和 $Z_n=X_1+X_2+\cdots+X_n$ 服从二项分布

$$\mathcal{B}(m_1+m_2+\cdots+m_n,p).$$

当然也可以按照例2.5.1的方法推导出上述结果, 但是从问题的背景出发得到的结果更加直观.

§2.5 随机向量函数的分布

B. 连续型随机向量函数的分布

设随机向量(X, Y)有联合密度$f(x, y)$, $U = u(x, y)$是二元函数, 则 $U = u(X, Y)$ 是随机变量. 于是可以研究 U 的概率密度的计算问题.

如果X, Y独立, 分别服从正态分布 $N(0, \sigma_1^2)$, $N(0, \sigma_2^2)$, 则称

$$R = \sqrt{X^2 + Y^2}$$

为脱靶量. 这是因为若将(X, Y)视为弹落点, 则R是弹落点到目标$(0, 0)$的距离.

§2.5 随机向量函数的分布

例2.5.2 设X, Y独立, 都服从标准正态分布$N(0,1)$, 求脱靶量$R=\sqrt{X^2+Y^2}$的概率密度.

解 (X,Y) 有联合密度

$$f(x,y)=\frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right).$$

R 在 $(0,\infty)$ 中取值. 定义 $D=\{(x,y)\,|\,x^2+y^2\leqslant r\}$. 对 $r>0$, 得到 R 的分布函数

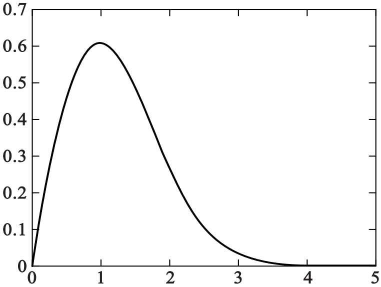
$$\begin{aligned}
F_R(r) &= P(\sqrt{X^2+Y^2}\leqslant r)\\
&= \iint_D \frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\int_0^r \mathrm{e}^{-z^2/2}z\,\mathrm{d}z \quad [\text{取}x=z\cos\theta, y=z\sin\theta]\\
&= \int_0^r \mathrm{e}^{-z^2/2}z\,\mathrm{d}z.
\end{aligned}$$

§2.5 随机向量函数的分布

$F_R(r)$ 连续，求导得到 R 的概率密度

$$f_R(r) = re^{-r^2/2}, \quad r > 0.$$

称为瑞利(Rayleigh)概率密度. 横轴是r，纵轴是$f_R(r)$.



瑞利概率密度

§2.5 随机向量函数的分布

值得指出，瑞利概率密度$f_R(r)$在$r=1$取最大值. 如果以原点为心画出若干宽度为2ε 的圆环，则子弹落在圆环

$$\{(x,y)|\, 1-\varepsilon<\sqrt{x^2+y^2}<1+\varepsilon\}$$

的概率较大. 这就解释了为什么优秀射击运动员在比赛时打出9或8环的机会较多，打出10 环或7, 6 环的机会较少.

§2.5 随机向量函数的分布

例2.5.3 设(X, Y)有联合密度$f(x, y)$, 则 $U = X + Y$ 有概率密度

$$f_U(u) = \int_{-\infty}^{\infty} f(x, u - x)\mathrm{d}x, \quad (2.5.1)$$

当 X, Y 独立时, $U = X + Y$ 有概率密度

$$f_U(u) = \int_{-\infty}^{\infty} f_X(x) f_Y(u - x)\mathrm{d}x, \quad (2.5.2)$$

§2.5 随机向量函数的分布

解 对$x > 0$, 利用$\mathrm{I}[x+y \leqslant u] = \mathrm{I}[y \leqslant u-x]$ 得到

$$\begin{aligned} F_U(u) &= P(U \leqslant u) = P(X+Y \leqslant u) \\ &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x,y)\mathrm{I}[x+y \leqslant u]\,\mathrm{d}y \Big) \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{u-x} f(x,y)\,\mathrm{d}y \Big) \mathrm{d}x. \end{aligned}$$

分布函数 $F_U(u)$ 是 u 的连续函数. 对 u 求导数, 并让求导数穿过第一个积分号, 得到U的概率密度(2.5.1). 当X, Y独立时, 由$f(x,y) = f_X(x)f_Y(y)$得到(2.5.2).

§2.5 随机向量函数的分布

例2.5.4 设(X, Y)有联合密度$f(x, y)$，则$V = X - Y$有概率密度

$$f_V(v) = \int_{-\infty}^{\infty} f(x, x - v)\mathrm{d}x.$$

特别当X，Y独立时，$V = X - Y$有概率密度

$$f_V(v) = \int_{-\infty}^{\infty} f_X(x) f_Y(x - v)\mathrm{d}x.$$

证明略.

例2.5.5 设X，Y独立，$X \sim \mathcal{E}(\lambda)$，$Y \sim \mathcal{E}(\mu)$. 求$U = X + Y$的概率密度.

§2.5 随机向量函数的分布

解 X, Y 分别有概率密度

$$f_X(x) = \lambda \mathrm{e}^{-\lambda x} \mathrm{I}[x > 0], \quad f_Y(y) = \mu \mathrm{e}^{-\mu y} \mathrm{I}[y > 0].$$

对于$u > 0$, 按公式(2.5.2)得到 U 的概率密度

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{\infty} f_X(x) f_Y(u-x)\,\mathrm{d}x \\ &= \int_0^{\infty} \lambda\mu \mathrm{e}^{-\lambda x} \mathrm{e}^{-\mu(u-x)} \mathrm{I}[x > 0]\mathrm{I}[u - x > 0]\,\mathrm{d}x \\ &= \lambda\mu \mathrm{e}^{-\mu u} \int_0^u \mathrm{e}^{-(\lambda-\mu)x}\,\mathrm{d}x \\ &= \begin{cases} \lambda\mu u \mathrm{e}^{-\mu u}, & \text{当}\lambda = \mu, \\ \frac{\lambda\mu}{\lambda-\mu}(\mathrm{e}^{-\mu u} - \mathrm{e}^{-\lambda u}), & \text{当}\lambda \neq \mu. \end{cases} \end{aligned}$$

§2.6 随机向量函数的密度

当X的概率密度$f(x)$在点x连续，则按微分法有$P(X=x)=f(x)\mathrm{d}x$. 当(X,Y)的联合密度$f(x,y)$在(x,y)连续，我们也用

$$P(X=x,Y=y)=g(x,y)\,\mathrm{d}x\mathrm{d}y$$

表示(X,Y)在点(x,y)有联合密度$g(x,y)$.

设D是R^2的子集，如果对D中每个点，都能在D中画出一个以该点为心的小圆，则称D是开集.

§2.6 随机向量函数的密度

定理2.6.1 如果平面的开集D使得$P((X,Y)\in D)=1$，且D中的连续函数$g(x,y)$使得

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

则

$$f(x,y)=g(x,y),\ (x,y)\in D$$

是(X,Y)的联合密度.

§2.6 随机向量函数的密度

和一元情况相似, 在应用定理2.6.1时, 要遵守以下约定:

(1) 只有在$A=B$时, 才能写$P(A)=P(B)$;

(2) 只有在$A=\bigcup_{i=1}^{n} A_i$,且$A_1, A_2, \cdots, A_n$作为集合互不相交时, 才能写

$$P(A)=\sum_{i=1}^{n} P(A_i).$$

§2.6 随机向量函数的密度

让我们再回忆微积分的知识：如果 $x = x(u, v)$, $y = y(u, v)$ 在平面的开集D内有连续的偏导数，并且雅可比行列式

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} \neq 0,$$

则有

$$\mathrm{d}x\mathrm{d}y = \left|\frac{\partial(x, y)}{\partial(u, v)}\right| \mathrm{d}u\mathrm{d}v = |J|\,\mathrm{d}u\mathrm{d}v,\ \ (u, v) \in D,$$

其中$|J|$是J的绝对值.

§2.6 随机向量函数的密度

例2.6.1 设 (X, Y) 有联合密度 $f(x, y)$, 计算最大值和最小值

$$U = \max(X, Y), \ \ V = \min(X, Y)$$

的联合密度.

解 对于开集$D = \{(u, v)|u > v\}$, 有

$$P((U, V) \in D) = P(X \neq Y) = \iint_{R^2} \mathrm{I}[x \neq y] f(x, y)\, \mathrm{d}x\mathrm{d}y = 1.$$

对于$(u, v) \in D$, 从

$$\begin{aligned} P(U = u, V = v) &= P(X = u, Y = v) + P(X = v, Y = u) \\ &= f(u, v)\, \mathrm{d}u\, \mathrm{d}v + f(v, u)\, \mathrm{d}v\, \mathrm{d}u \\ &= [f(u, v) + f(v, u)]\, \mathrm{d}u\, \mathrm{d}v \end{aligned}$$

知道(U, V)的联合密度是

$$g(u, v) = f(u, v) + f(v, u), \ \ u > v.$$

§2.6 随机向量函数的密度

例2.6.1 设 (X, Y) 有联合密度 $f(x, y)$, 计算最大值和最小值

$$U = \max(X, Y), \ \ V = \min(X, Y)$$

的联合密度.

解 对于开集$D = \{(u, v) | u > v\}$, 有

$$P((U, V) \in D) = P(X \neq Y) = \iint_{R^2} \mathrm{I}[x \neq y] f(x, y) \, \mathrm{d}x\mathrm{d}y = 1.$$

对于$(u, v) \in D$, 从

$$\begin{aligned} P(U = u, V = v) &= P(X = u, Y = v) + P(X = v, Y = u) \\ &= f(u, v) \, \mathrm{d}u \, \mathrm{d}v + f(v, u) \, \mathrm{d}v \, \mathrm{d}u \\ &= [f(u, v) + f(v, u)] \, \mathrm{d}u \, \mathrm{d}v \end{aligned}$$

知道(U, V)的联合密度是

$$g(u, v) = f(u, v) + f(v, u), \ \ u > v.$$

§2.6 随机向量函数的密度

例2.6.2 设 (X, Y) 有联合密度 $f(x, y)$, (U, V) 由线性变换

$$U = 2X - Y, \quad V = 2X + 3Y$$

决定, 求(U, V)的联合密度.

解 从$u = 2x - y$, $v = 2x + 3y$ 解出

$$x = (3u + v)/8, \quad y = (-u + v)/4,$$

并且

$$J^{-1} = \frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} 2 & -1 \\ 2 & 3 \end{vmatrix} = 8, \quad J = \frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{8}.$$

§2.6 随机向量函数的密度

例2.6.2 设 (X, Y) 有联合密度 $f(x, y)$, (U, V) 由线性变换

$$U = 2X - Y, \quad V = 2X + 3Y$$

决定, 求(U, V)的联合密度.

解 从$u = 2x - y$, $v = 2x + 3y$ 解出

$$x = (3u + v)/8, \quad y = (-u + v)/4,$$

并且

$$J^{-1} = \frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} 2 & -1 \\ 2 & 3 \end{vmatrix} = 8, \quad J = \frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{8}.$$

§2.6 随机向量函数的密度

对于(u,v), 从

$$
\begin{aligned}
P(U=u,V=v) &= P\Big(2X-Y=u, 2X+3Y=v\Big) \\
&= P\Big(X=\frac{3u+v}{8}, Y=\frac{v-u}{4}\Big) \\
&= f\Big(\frac{3u+v}{8}, \frac{v-u}{4}\Big)|J|\,\mathrm{d}u\mathrm{d}v
\end{aligned}
$$

得到(U,V)的联合密度

$$
g(u,v)=\frac{1}{8}f\Big(\frac{3u+v}{8}, \frac{v-u}{4}\Big).
$$

§2.6 随机向量函数的密度

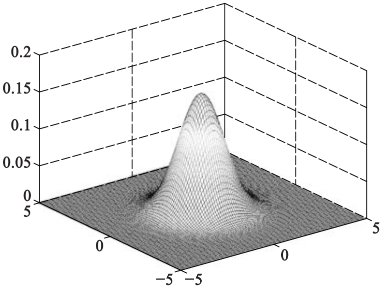
二维正态分布

如果Y_1, Y_2独立, 都服从标准正态分布 $N(0,1)$, 则 $\boldsymbol{Y} = (Y_1, Y_2)$ 有联合密度

$$\varphi(y_1, y_2) = \frac{1}{2\pi} \exp\left(-\frac{y_1^2 + y_2^2}{2}\right).$$

这时称 $\boldsymbol{Y} = (Y_1, Y_2)$ 服从二维标准正态分布, 记做 $\boldsymbol{Y} \sim N(\boldsymbol{0}, \boldsymbol{I})$.

§2.6 随机向量函数的密度



二维标准正态密度$\varphi(y_1, y_2)$

下面从二维标准正态分布引入二维正态分布.

§2.6 随机向量函数的密度

设$\boldsymbol{Y}=(Y_1,Y_2)$服从二维标准正态分布$N(\boldsymbol{0},\boldsymbol{I})$, $ad-bc\neq 0$. 定义

$$\begin{cases} X_1 = aY_1 + bY_2 + \mu_1, \\ X_2 = cY_1 + dY_2 + \mu_2, \end{cases}$$

则称(X_1,X_2)服从的分布为二维正态分布.

§2.6 随机向量函数的密度

引入

$$\sigma_1 = \sqrt{a^2 + b^2}, \ \sigma_2 = \sqrt{c^2 + d^2}, \quad \rho = (ac + bd)/(\sigma_1\sigma_2).$$

则可以把$\boldsymbol{X} = (X_1, X_2)$的联合密度写成(略去推导)

$$\begin{aligned} f(x_1, x_2) \quad &= \quad \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} \right.\right. \\ &\quad \left.\left. - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right] \right\}. \qquad (2.6.1) \end{aligned}$$

因为(2.6.1)中只有5个参数$\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho$, 所以又称 (X_1, X_2) 服从参数为 $(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$ 的正态分布, 记做

$$(X_1, X_2) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho).$$

§2.6 随机向量函数的密度

定理2.6.1　如果(X_1, X_2)有联合密度(2.6.1),则

(1) $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$;
(2) X_1, X_2独立的充分必要条件是$\rho = 0$;
(3) 当$a_1 a_4 - a_3 a_2 \neq 0$, 随机向量(Z_1, Z_2)服从二维正态分布, 其中

$$\begin{cases} Z_1 = a_1 X_1 + a_2 X_2 + c_1, \\ Z_2 = a_3 X_1 + a_4 X_2 + c_2; \end{cases}$$

(4) 线性组合$Z_1 = a_1 X_1 + a_2 X_2 + c_1$服从正态分布;
(5) 若Z_1, Z_2相互独立, 都服从正态分布, 则(Z_1, Z_2)服从二维正态分布.

§4.6　二维正态分布

用定理2.6.1容易验证以下三个结论:

- 如果$\boldsymbol{Y}=(Y_1,Y_2)$服从二维标准正态分布$N(\boldsymbol{0},\boldsymbol{I})$,
$$X_1=Y_1+bY_2,\ X_2=-bY_1+Y_2,$$
则X_1, X_2独立, 都服从正态分布$N(0,1+b^2)$;
- 如果$\boldsymbol{Y}=(Y_1,Y_2)$服从二维标准正态分布$N(\boldsymbol{0},\boldsymbol{I})$, a, b是不全为0的常数, 则线性组合$X=aY_1+bY_2+c$服从正态分布;
- 如果$X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$ 独立, 且X,Y独立, 常数$a\neq 0$, 则线性组合
$$U=aX+bY+c$$
服从正态分布.

§2.6 随机向量函数的密度

- 设想一粒花粉在水面由于受到水分子的碰撞而作布朗运动. 设想在水面建立一个直角坐标系, 使得花粉运动的起点是(μ_1, μ_2). 用(X, Y)表示花粉在t时刻的坐标. 由于花粉的运动是各向同性的, 所以(X, Y) 在(μ_1, μ_2) 周围取值的概率较大, 在离开 (μ_1, μ_2) 较远的地方取值的概率较小. 二维正态概率密度正好描述这一现象.
- 设想运动员的打靶, 用(μ_1, μ_2)表示靶心的坐标, 用(X, Y) 表示弹落点. 则(X, Y) 落在(μ_1, μ_2)附近的概率较大, 落在较远的地方的概率较小.二维正态联合密度也正好描述这一现象.

§2.7　条件分布

设 A, B是事件, $P(A)>0$. 已知 A发生的条件下, B 的条件概率是

$$P(B|A)=\frac{P(AB)}{P(A)}.$$

以后称 $P(B|A)$是 $B|A$ 的概率. 完全类似地, 如果随机变量 X 的取值是 $x_1, x_2, \cdots$, 则称

$$h_i=P(X=x_i|A),\ i=1,2,\cdots$$

为$X|A$的概率分布.

§2.7 条件分布

例2.7.1 设(X, Y)是离散型随机向量，有联合分布

$$p_{ij} = P(X = x_i, Y = y_j) > 0, \quad i, j = 1, 2, \cdots .$$

(a) 对确定的j，计算$X|\{Y = y_j\}$的分布;

(b) 对确定的i，计算$Y|\{X = x_i\}$的分布.

§2.7 条件分布

解 X, Y 分别有边缘分布

$$\begin{aligned} p_i &= P(X = x_i) = \sum_{j=1}^{\infty} p_{ij}, \ i = 1, 2, \cdots, \\ q_j &= P(Y = y_j) = \sum_{i=1}^{\infty} p_{ij}, \ j = 1, 2, \cdots. \end{aligned}$$

根据条件概率公式得到 $X|\{Y = y_j\}$ 的概率分布

$$P(X = x_i|Y = y_j) = \frac{p_{ij}}{q_j}, \quad i = 1, 2, \cdots.$$

同理得到 $Y|\{X = x_i\}$ 的概率分布

$$P(Y = y_j|X = x_i) = \frac{p_{ij}}{p_i}, \ j = 1, 2, \cdots.$$

§2.7 条件分布

例2.7.2 甲向一个目标独立重复射击，用S_n表示第n次击中目标时的射击次数. 如果甲每次击中目标的概率是$p = 1 - q$, 计算

(a) (S_1, S_2)的联合分布;

(b) S_2的概率分布;

(c) $S_1|\{S_2 = j\}$的概率分布.

§2.7 条件分布

解 (a) 根据题意得到 (S_1, S_2) 的联合分布

$$P(S_1 = i, S_2 = j) = q^{i-1} p q^{j-i-1} p = p^2 q^{j-2},\ j > i \geqslant 1.$$

(b) S_2的边缘密度是

$$\begin{aligned} P(S_2 = j) &= \sum_{i=1}^{j-1} P(S_1 = i, S_2 = j) \\ &= \sum_{i=1}^{j-1} p^2 q^{j-2} = (j-1) p^2 q^{j-2},\quad j = 2, 3, \cdots. \end{aligned}$$

§2.7 条件分布

解 (a) 根据题意得到 (S_1, S_2) 的联合分布

$$P(S_1 = i, S_2 = j) = q^{i-1}pq^{j-i-1}p = p^2q^{j-2},\ j > i \geqslant 1.$$

(b) S_2的边缘密度是

$$\begin{aligned} P(S_2 = j) &= \sum_{i=1}^{j-1} P(S_1 = i, S_2 = j) \\ &= \sum_{i=1}^{j-1} p^2q^{j-2} = (j-1)p^2q^{j-2}, \quad j = 2, 3, \cdots. \end{aligned}$$

§2.7 条件分布

(c) 从(a)和(b)的结论得到$S_1|\{S_2=j\}$的概率分布

$$P(S_1=i|S_2=j) = \frac{P(S_1=i,S_2=j)}{P(S_2=j)} = \frac{p^2q^{j-2}}{(j-1)p^2q^{j-2}} = \frac{1}{j-1},\ 1\leqslant i<j. \qquad (2.7.1)$$

公式(2.7.1)说明$S_1|\{S_2=j\}$ 在 $\{1,2,\cdots,j-1\}$中等可能地取值. 也就是说已知 $S_2=j$时, S_1 在 $\{1,2,\cdots,j-1\}$ 中的取值是等可能的.

§2.7 条件分布

近30年来，随着生活水平的日益提高，北京人普遍认为中小学生的平均身高有明显的增长. 为了证实这一情况，需要对中小学生的身高现状进行抽样调查. 因为身高和年龄密切相关，所以抽样调查时必须同时考虑年龄的因素. 在北京随机选取一个男生，用X和Y 分别表示他的身高和年龄时，得到随机向量(X, Y). 因为对任何x, y，理论上讲有$P(X = x, Y = y) = 0$，所以(X, Y)的联合分布函数$F(x, y)$ 应当是连续函数. 假设$F(x, y)$有连续的偏导数，则(X, Y)是连续型随机向量,有联合密度

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

这时Y有边缘密度

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

§2.7 条件分布

如果只调查年龄为15周岁的男生的身高情况，则相当于在条件 $Y = 15$ 下，研究 X 的分布和平均身高. 所以研究 $X|\{Y = 15\}$ 的概率分布是有明确意义的实际问题. 同理，对于任何使得 $f_Y(y) > 0$ 的 y，研究 $X|\{Y = y\}$ 的概率分布都是有意义的问题.

对于使得 $f_Y(y) > 0$ 的 y，例如 $y = 15$，形式的推导给出

$$
\begin{aligned}
P(X = x|Y = y) &= \frac{P(X = x, Y = y)}{P(Y = y)} \\
&= \frac{f(x, y)\,\mathrm{d}x\mathrm{d}y}{f_Y(y)\,\mathrm{d}y} \\
&= \frac{f(x, y)}{f_Y(y)}\,\mathrm{d}x.
\end{aligned}
$$

§2.7 条件分布

根据微分法知道, 已知$Y=y$的条件下, X 有条件密度

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}. \qquad (2.7.2)$$

以后称$f_{X|Y}(x|y)$为$X|\{Y=y\}$的概率密度.

- 当$y=15$, 条件密度$f_{X|Y}(x|y)$就是年龄为15 周岁的男生身高的概率密度.

§2.7 条件分布

如果用 $F_{X|Y}(x|y) = P(X \leqslant x|Y = y)$ 表示 $X|\{Y = y\}$ 的分布函数，则根据概率密度和分布函数的关系知道

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(s|y)\,\mathrm{d}s = \int_{-\infty}^{x} \frac{f(s,y)}{f_Y(y)}\,\mathrm{d}s.$$

以后称$F_{X|Y}(x|y)$为$X|\{Y = y\}$ 的分布函数.

- 当$y = 15$，条件分布函数$F_{X|Y}(x|y) = F_{X|Y}(x|15)$就是年龄为15周岁的男生身高的分布函数.

§2.7　条件分布

定义2.7.1　设随机向量(X,Y)有联合密度$f(x,y)$, Y有边缘密度$f_Y(y)$. 如果在y处$f_Y(y)>0$, 则称

$$F_{X|Y}(x|y)=P(X\leqslant x|Y=y)=\int_{-\infty}^{x}\frac{f(s,y)}{f_Y(y)}\,\mathrm{d}s,\quad x\in R$$

为$X|\{Y=y\}$的分布函数, 简称为条件分布函数. 称

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)},\quad x\in R \qquad (2.7.3)$$

为$X|\{Y=y\}$的概率密度, 简称为条件密度.

§2.7　条件分布

根据定义2.7.1可以得到条件密度和条件分布函数的关系如下：如果y使得$f_Y(y) > 0$，则

$$F_{X|Y}(x|y) = P(X \leqslant x|Y = y) = \int_{-\infty}^{x} f_{X|Y}(s|y)\,\mathrm{d}s,\ x \in R.$$

容易看出，X，Y独立的充分必要条件是

$$F_{X|Y}(x|y) = F_X(x) \quad \text{或} \quad f_{X|Y}(x|y) = f_X(x)$$

之一对所有的x, y成立.

§2.7　条件分布

例2.7.3　假设北京市男生年龄的概率密度是 $h(y)$,年龄为y的男生的身高的概率密度是 $g(x,y)$. 计算北京市男生的身高和年龄的联合密度.

解　用X, Y分别表示男生的身高和年龄. 根据题意, 年龄Y有概率密度 $f_Y(y)=h(y)$, 年龄为 y 的男生身高 $X|\{Y=y\}$ 有条件密度$f_{X|Y}(x|y)=g(x,y)$. 从公式(2.7.3)得到(X,Y) 的联合密度

$$f(x,y)=f_Y(y)f_{X|Y}(x|y)=h(y)g(x,y).$$

从公式(2.7.3)看出: 对于固定的y, 条件密度$f_{X|Y}(x|y)$ 是x的函数且和联合密度$f(x,y)$只相差常数因子$1/f_Y(y)$.

以上性质可以加深我们对条件密度的几何理解, 并帮助我们方便地计算条件密度. 看下面的例子.

§2.7 条件分布

例2.7.3 假设北京市男生年龄的概率密度是 $h(y)$,年龄为y的男生的身高的概率密度是 $g(x,y)$. 计算北京市男生的身高和年龄的联合密度.

解 用X, Y分别表示男生的身高和年龄. 根据题意, 年龄Y有概率密度 $f_Y(y)=h(y)$, 年龄为 y 的男生身高 $X|\{Y=y\}$ 有条件密度$f_{X|Y}(x|y)=g(x,y)$. 从公式(2.7.3)得到(X,Y) 的联合密度

$$f(x,y)=f_Y(y)f_{X|Y}(x|y)=h(y)g(x,y).$$

从公式(2.7.3)看出: 对于固定的y, 条件密度$f_{X|Y}(x|y)$ 是x的函数且和联合密度$f(x,y)$只相差常数因子$1/f_Y(y)$.

以上性质可以加深我们对条件密度的几何理解, 并帮助我们方便地计算条件密度. 看下面的例子.

§2.7 条件分布

例2.7.3 假设北京市男生年龄的概率密度是 $h(y)$,年龄为y的男生的身高的概率密度是 $g(x,y)$. 计算北京市男生的身高和年龄的联合密度.

解 用X, Y分别表示男生的身高和年龄. 根据题意, 年龄Y有概率密度 $f_Y(y)=h(y)$, 年龄为 y 的男生身高 $X|\{Y=y\}$ 有条件密度$f_{X|Y}(x|y)=g(x,y)$. 从公式(2.7.3)得到(X,Y) 的联合密度

$$f(x,y)=f_Y(y)f_{X|Y}(x|y)=h(y)g(x,y).$$

从公式(2.7.3)看出: 对于固定的y, 条件密度$f_{X|Y}(x|y)$ 是x的函数且和联合密度$f(x,y)$只相差常数因子$1/f_Y(y)$.

以上性质可以加深我们对条件密度的几何理解, 并帮助我们方便地计算条件密度. 看下面的例子.

§2.7 条件分布

例2.7.4 设 (X, Y) 在单位圆 $D=\{(x, y)|x^2+y^2 \leqslant 1\}$ 内均匀分布. 计算条件密度 $f_{X|Y}(x|y)$.

解 对 $|y|<1$, 设 $C_y=\sqrt{1-y^2}$, 则 (X, Y) 的联合密度是

$$f(x, y)=\frac{1}{\pi}\mathrm{I}[x^2+y^2 \leqslant 1]=\frac{1}{\pi}\mathrm{I}[|x| \leqslant C_y].$$

已知 $Y=y$ 时, 因为当 x 在 $[-C_y, C_y]$ 中变化时 $f(x, y)$ 是常数, 所以 $f_{X|Y}(x|y)$ 在 $[-C_y, C_y]$ 中也是常数. 说明 $X|\{Y=y\}$ 在 $[-C_y, C_y]$ 中均匀分布. 于是, 对 $y \in(-1,1)$,

$$f_{X|Y}(x|y)=\frac{1}{2C_y}=\frac{1}{2\sqrt{1-y^2}}, \quad x \in[-C_y, C_y].$$

§2.7 条件分布

例2.7.4 设 (X, Y) 在单位圆 $D = \{(x, y)|x^2 + y^2 \leqslant 1\}$ 内均匀分布. 计算条件密度 $f_{X|Y}(x|y)$.

解 对$|y| < 1$, 设 $C_y = \sqrt{1 - y^2}$, 则 (X, Y) 的联合密度是

$$f(x, y) = \frac{1}{\pi}\mathrm{I}[x^2 + y^2 \leqslant 1] = \frac{1}{\pi}\mathrm{I}[|x| \leqslant C_y].$$

已知$Y = y$时, 因为当x在$[-C_y, C_y]$中变化时$f(x, y)$是常数, 所以$f_{X|Y}(x|y)$ 在 $[-C_y, C_y]$ 中也是常数. 说明 $X|\{Y = y\}$ 在 $[-C_y, C_y]$ 中均匀分布. 于是, 对 $y \in (-1, 1)$,

$$f_{X|Y}(x|y) = \frac{1}{2C_y} = \frac{1}{2\sqrt{1 - y^2}}, \quad x \in [-C_y, C_y].$$

§2.7 条件分布

例2.7.5 设炮击的目标是(μ_1, μ_2), 弹落点的坐标(X, Y) 服从正态分布 $N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$, 有概率密度

$$
\begin{aligned}
f(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Big\{ -\frac{1}{2(1-\rho^2)}\Big[\frac{(x-\mu_1)^2}{\sigma_1^2} \\
&\qquad -\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\Big]\Big\}.
\end{aligned}
$$

已知弹落点的纵坐标是y时, 计算弹落点横坐标的概率密度.

解 对确定的y, 需要计算$X|\{Y=y\}$的概率密度$f_{X|Y}(x|y)$. 定义

$$\mu_y = \mu_1 + \rho(\sigma_1/\sigma_2)(y-\mu_2),\ \sigma_y^2 = (1-\rho^2)\sigma_1^2. \quad (*)$$

§2.7 条件分布

例2.7.5 设炮击的目标是(μ_1,μ_2)，弹落点的坐标(X,Y)服从正态分布$N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$，有概率密度

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}\right.\right. \\ &\qquad \left.\left.-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}. \end{aligned}$$

已知弹落点的纵坐标是y时，计算弹落点横坐标的概率密度.

解 对确定的y，需要计算$X|\{Y=y\}$的概率密度$f_{X|Y}(x|y)$. 定义

$$\mu_y=\mu_1+\rho(\sigma_1/\sigma_2)(y-\mu_2),\ \sigma_y^2=(1-\rho^2)\sigma_1^2. \quad (*)$$

§2.7 条件分布

用$A(x) \propto B(x)$表示函数$A(x)$和$B(x)$相差一个常数因子. 对于确定的y, 作为x的函数, 有

$$\begin{aligned}
f_{X|Y}(x|y) \propto\ & f(x,y) \\
\propto\ & \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right]\right\} \\
\propto\ & \exp\left\{-\frac{1}{2(1-\rho^2)\sigma_1^2}\left[(x-\mu_1)^2-2\rho(\sigma_1/\sigma_2)(x-\mu_1)(y-\mu_2)\right]\right\} \\
\propto\ & \exp\left\{-\frac{1}{2\sigma_y^2}\left[x-\mu_1-\rho(\sigma_1/\sigma_2)(y-\mu_2)\right]^2\right\} \\
\propto\ & \frac{1}{\sqrt{2\pi}\,\sigma_y}\exp\left[-\frac{(x-\mu_y)^2}{2\sigma_y^2}\right].
\end{aligned}$$

说明 $X|\{Y=y\}$ 服从正态分布 $N(\mu_y,\sigma_y^2)$.

§2.7　条件分布

从例2.7.5可以总结出以下定理.

定理2.7.1　如果(X,Y)服从二维正态分布

$$N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho),$$

则 $X|\{Y=y\}$ 服从正态分布 $N(\mu_y,\sigma_y^2)$, 其中 μ_y, σ_y^2 如 $(*)$.