第二章随机变量及其分布

2.3 多维分布与边际分布

2.3.1 多维分布

在实际应用中，经常需要对所考虑的问题用多个变量来描述. 我们把多个随机变量放在一起组成向量，称为多维随机变量或者随机向量.

从一副扑克牌中抽牌时, 可以用纸牌的花色和数字来说明其特征.

↑Example

↓Example

考虑一个打靶的试验. 在靶面上取定一个直角坐标系. 则命中的位置可由其坐标 (X, Y) 来刻划. X,Y 都是随机变量.

↑Example

↓Example

Definition

设 $X=(X_1,...,X_n)$. 如果每个 X_i 都是一个随机变量, $i=1,\cdots,n$, 则称 X 为 n 维随机变量或者随机向量.

我们可以按照对常用一维随机变量的分类把常用的随机向量分为**离散型**、**连续型**以及**其他类型**.

如果每一个 X_i 都是一个离散型随机变量，$i = 1, ..., n$，则称 $X = (X_1, \ldots, X_n)$ 为一 n 维离散随机变量．设 X_i 的所有可能取值（有限或可数个）为 $\{a_{i1}, a_{i2}, \cdots\}$，$i = 1, \ldots, n$，则称

Definition

$$p(j_1, \cdots, j_n) = P(X_1 = a_{1j_1}, \ldots, X_n = a_{nj_n}),\ j_1, ..., j_n = 1, 2, ... \tag{2.1}$$

为 n 维随机变量 X 的概率函数．

容易证明概率函数具有下列性质:

(1) $p(j_1, \ldots, j_n) \geq 0, \quad j_i = 1, 2, \cdots,\ i = 1, 2, \ldots, n;$

(2) $\sum\limits_{j_1, \cdots, j_n} p(j_1, \ldots, j_n) = 1.$

$\overline{\uparrow}$Example

设 $A_1,\cdots,A_n$ 为某一实验下的完备事件群，即 $A_1,\cdots,A_n$ 两两互斥且和为 Ω。记 $p_k=P(A_k)(k=1,\ldots,n)$，则 $p_k\geq 0, p_1+\cdots+p_n=1$。现将实验独立的重复作 N 次，分别用 X_i 表示事件 A_i 出现的次数 $(i=1,\cdots,n)$。则 $X=(X_1,\ldots,X_n)$ 为一离散型随机向量，试求 X 的概率函数。此分布律称为多项分布，记为 $M(N;p_1,\ldots,p_n)$.

$\underline{\downarrow}$Example

我们来看一下 X_i 的分布：此时我们把试验结果分为两类，A_i 和 $\bar{A}_i$，则显然就是一个 N 重贝努里试验，因此

$$P(X_i = k_i) = \binom{N}{k_i} p_i^{k_i} (1-p_i)^{N-k_i}, \quad k_i = 1, \cdots, N.$$

类似我们也可以找出 $(X_i, X_j)(i \neq j)$ 的联合分布律，即为 $M(N, p_i, p_j, 1-p_i-p_j)$.

我们具体来看一下二维离散分布．设二维离散型随机变量 (X,Y) 的所有可能取值为 $\{(x_i,y_j): i=1,...,n, j=1,2,...,m\}$. 这里 n,m 为有限数或者无穷．我们经常以列联表的形式来表示二维离散型随机变量的概率分布．记

$$p_{ij}=P(X=x_i,Y=y_j),\ i=1,...,n,j=1,...,m.$$

则 (X,Y) 的概率函数可以用下表表示:

$X \backslash Y$	y_1	x_2	$\cdots$	y_m	行和
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}	$p_{1\cdot}$
x_2	p_{12}	p_{22}	$\cdots$	p_{2m}	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\vdots$	p_{nm}	$p_{n\cdot}$
列和	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot m}$	1

↑Example

从一个包含五个黑球, 六个白球和七个红球的罐子里抽取四个球. 令 X 是抽到白球的数目, Y 是抽到红球的数目. 则二维随机变量 (X,Y) 的概率函数为

$$p(x,y)=\frac{\binom{6}{x}\binom{7}{y}\binom{5}{4-x-y}}{\binom{18}{4}},\ 0\leq x+y\leq 4. \tag{2.2}$$

↓Example

以列联表表示, 即为

$Y \backslash X$	0	1	2	3	4	行和
0	$\frac{1}{612}$	$\frac{1}{51}$	$\frac{5}{102}$	$\frac{5}{153}$	$\frac{1}{204}$	$\frac{11}{102}$
1	$\frac{7}{306}$	$\frac{7}{51}$	$\frac{35}{204}$	$\frac{7}{153}$		$\frac{77}{204}$
2	$\frac{7}{102}$	$\frac{7}{34}$	$\frac{7}{68}$			$\frac{7}{17}$
3	$\frac{35}{612}$	$\frac{7}{102}$				$\frac{77}{612}$
4	$\frac{7}{612}$					$\frac{7}{612}$
列和	$\frac{99}{612}$	$\frac{22}{51}$	$\frac{11}{34}$	$\frac{4}{51}$	$\frac{1}{204}$	1

类似于一维连续型随机变量, 连续型随机向量的也是由密度函数来刻画的.

Definition

称 $X=(X_1,\ldots,X_n)$ 为 n 维连续型随机变量，如果存在 $\mathbb{R}^n$ 上的非负函数 $f(x_1,\ \ldots,\ x_n)$，使得对任意的 $-\infty < a_1 \le b_1 < +\infty, ..., -\infty < a_n \le b_n < +\infty$，有

$$P(a_1 \le X_1 \le b_1, ..., a_n \le X_n \le b_n) = \int_{a_n}^{b_n} ... \int_{a_1}^{b_1} f(x_1,\ldots,x_n)dx_1 \cdots dx_n,$$

则称 f 为 X 的概率密度函数.

称 $X=(X_1,\ldots,X_n)$ 为 n 维连续型随机变量，如果存在 $\mathbb{R}^n$ 上的非负函数 $f(x_1，\ldots，x_n)$，使得对任意的 $-\infty<x_1<+\infty,...,-\infty<x_n<+\infty$，有

$$F(x_1,...,x_n)=\int_{-\infty}^{x_n}...\int_{-\infty}^{x_1}f(t_1,\ldots,t_n)dt_1\cdots dt_n,$$

则称 f 为 X 的概率密度函数.

Definition

对 n 维随机变量我们也有分布函数的概念.

设 $X = (X_1, \dots, X_n)$ 为 n 维随机变量. 对任意的 $(x_1, \dots, x_n) \in \mathbb{R}^n$, 称

$$F(x_1, \dots, x_n) = P(X_1 \le x_1, \dots, X_n \le x_n) \qquad (2.3)$$

为 n 维随机变量 X 的 (联合) 分布函数.

Definition

可以验证分布函数 $F(x_1, \dots, x_n)$ 具有下述性质:

(1) $F(x_1, \cdots, x_n)$ 对每个变元单调非降;

(2) 对任意的 $1 \le j \le n$ 有, $\lim\limits_{x_j \to -\infty} F(x_1, \cdots, x_n) = 0$;

(3) $\lim\limits_{x_1 \to \infty, \cdots, x_n \to \infty} F(x_1, \cdots, x_n) = 1$.

对 n 维连续型随机变量, 从密度的定义我们有,

$$F(x_1,\ldots,x_n)=\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_1}f(x_1,...,x_n)dx_1...dx_n.$$

对高维离散型随机变量, 一般我们不使用分布函数.

↑Example

考虑二维随机变量 $X=(X_1,X_2)$, 其概率密度函数为

$$f(x_1,x_2)=\begin{cases}1/[(b-a)(d-c)] & \text{当 } a\le x_1\le b,\ c\le x_2\le d,\\ 0 & \text{其它.}\end{cases}$$

称此概率密度为 $[a,b]\times[c,d]$ 上的均匀分布.

↓Example

↑Example

设 (X, Y) 的概率密度函数有形式

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

其中 $-\infty < a, b < \infty$, $0 < \sigma_1, \sigma_2 < \infty$, $-1 \leq \rho \leq 1$. 称 (X, Y) 服从参数为 $a, b, \sigma_1, \sigma_2, \rho$ 的二元正态分布，记为 $N(a, b, \sigma_1^2, \sigma_2^2, \rho)$.

↓Example

2.3.2 边缘分布

设 $(X_1, ..., X_n) \sim F$ 已知. 令 $(i_1, ..., i_m) \subset (1, ..., n)$, 则 $X_{i_1}, ..., X_{i_m}$ 的分布称为 $X_1, ..., X_n$ 或 F 的一个m维边缘分布. **如何得到该分布?**

我们先考虑离散型随机向量. 设二维离散随机变量 (X, Y) 的所有可能取值为 $\{(x_i, y_j) : i = 1, ..., n, j = 1, 2, ..., m.\}$(这里 n, m 为有限数或者无穷), 则 (X, Y) 的联合分布律为

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

以列联表的形式表示就是

$Y \backslash X$	x_1	x_2	$\cdots$	x_n	行和
y_1	p_{11}	p_{21}	$\cdots$	p_{n1}	$p_{\cdot 1}$
y_2	p_{12}	p_{22}	$\cdots$	p_{n2}	$p_{\cdot 2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m	p_{1m}	p_{2m}	$\vdots$	p_{nm}	$p_{\cdot m}$
列和	$p_{1\cdot}$	$p_{2\cdot}$	$\cdots$	$p_{n\cdot}$	1

从上述列联表我们可以计算随机变量 X 和 Y 的分布. 固定某个 x_i. 因为 Y 在使得 $X = x_i$ 的那些样本点上必取值为 $y_1, ..., y_m$ 中之一, 故有

$$p_X(x_i) = P(X = x_i) = \sum_j^m P(X = x_i, Y = y_j) = \sum_j^m p_{ij} = p_{i\cdot}, \quad i = 1, 2, \cdots n.$$

所以上述列联表的行和所表示的正是 X 的分布. 因为这个分布是从 X 和 Y 的联合分布推导出来的, 我们称其为 X 的边缘分布.

类似可以得到 Y 的边缘分布律

$$p_Y(y_j) = P(Y = y_j) = \sum_{i}^{n} p_{ij} = p_{\cdot j},\ j = 1, 2, \cdots m.$$

它是上述列联表的列和.

n 维场合:

类似地，可对 n $(n > 2)$ 维的随机变量定义边缘分布. 设 $X_1, ..., X_n$ 为 n 维随机变量，其概率分布 F 已知. 令 $i_1 < \cdots < i_m$ 为 $1, ..., n$ 的任一子集，则 $X_{i_1}, ..., X_{i_m}$ 的概率函数为

$$
\begin{aligned}
& p_{i_1 \ldots i_m}(j_{i_1}, ..., j_{i_m}) = P(X_{i_1} = a_{i_1 j_{i_1}}, ..., X_{i_m} = a_{i_m j_{i_m}}) \\
= \;& P(X_{i_1} = a_{i_1 j_{i_1}}, ..., X_{i_m} = a_{i_m j_{i_m}}) \\
= \;& \sum_{j_{i_{m+1}}, ..., j_{i_n}} P(X_1 = a_{1 j_1}, ..., X_{i_1} = a_{i_1 j_{i_1}}, ..., X_{i_m} = a_{i_m j_{i_m}}, \\
& X_{i_{m+1}} = a_{i_{m+1} j_{i_{m+1}}}, ..., X_n = a_{n j_n}) \\
= \;& \sum_{\text{除} j_{i_1}, ..., j_{i_m} \text{的所有}} p(j_1, ..., j_n).
\end{aligned}
$$

其中和是对除 $X_{i_1}, ..., X_{i_m}$ 之外的所有变量来求和.

袋中有 5 张外形相同的卡片，其中 3 张写上数字"0"，另 2 张写上"1"。现从袋中任取两张卡片，分别以 ξ,η 表示第一张和第二张卡片上的数字，试求分别在有放回和不放回两种情形下 (ξ,η) 的联合分布律及边际分布律.

↑Example ↓Example

解：简单计算得到

这个例子说明**边际分布律不能决定联合分布律**。

现考虑连续型随机向量的边缘分布. 先考虑二维的情形. 设 (X,Y) 有概率密度函数 $f(x,y)$. 则

$$
\begin{aligned}
P(x_1 \le X \le x_2) &= P(x_1 \le X \le x_2, -\infty < Y < +\infty) \\
&= \int_{-\infty}^{+\infty} \int_{x_1}^{x_2} f(u,v)dudv \\
&= \int_{x_1}^{x_2} f_X(u)du, \qquad (2.4)
\end{aligned}
$$

其中

$$
f_X(u) = \int_{-\infty}^{+\infty} f(u,v)dv. \qquad (2.5)
$$

从 (2.4) 我们可以看出, X 的边缘密度函数即为 (2.5). 类似地, Y 的边缘密度函数为

$$
f_Y(u) = \int_{-\infty}^{+\infty} f(u,v)du. \qquad (2.6)
$$

当 $n>2$ 时, 令 $f(x_1,...,x_n)$ 为 n 维连续型随机变量 $(X_1,...,X_n)$ 的概率密度函数. 设 $(i_1,\cdots,i_m)$ 为 $(1,2,...,n)$ 的一个子集. 则同上可证, 则 $(X_{i_1},...,X_{i_m})$ 的概率密度函数是联合密度函数 $f(x_1,...,x_n)$ 对除 $X_{i_1},...,X_{i_m}$ 之外的所有变量求积分.

设 (X_1,X_2) 服从 $N(a,b,\sigma_1^2,\sigma_2^2,\rho)$. 则可证明 X_1 的边缘分布为 $N(a,\sigma_1^2)$, X_2 的边缘分布为 $N(b,\sigma_2^2)$.

↑Example
↓Example

例2.3.2说明了虽然 n 维随机变量 $X=(X_1,...,X_n)$ 的分布可以唯一决定其所有的边缘分布，但边缘分布不足以决定 X 的联合分布.

↑Example

考虑两个概率密度函数

$$\begin{aligned} p(x,y) &= x+y, \qquad 0<x,y<1 \\ q(x,y) &= (x+\frac{1}{2})(y+\frac{1}{2}), \qquad 0<x,y<1 \end{aligned}$$

试求边际概率密度。

↓Example

解：易得所求边际概率密度都是如下形式

$$f(t)=t+\frac{1}{2}, \quad 0<t<1.$$

说明**边际概率密度不能决定联合概率密度**。

设 (X,Y) 的联合概率密度有形式 $(\forall (x,y)\in \mathbb{R}^2)$

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

其中 $-\infty < a,b < \infty; 0 < \sigma_1,\sigma_2 < \infty; -1 \le \rho \le 1$. 则称 (X,Y) 服从参数为 $a,b,\sigma_1,\sigma_2,\rho$ 的二元正态分布，记为 $N(a,b,\sigma_1^2,\sigma_2^2,\rho)$. 试计算 X 和 Y 的边际概率密度。

解: