目 录

第二章　随机变量及其分布

§2.4　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.

例 2.4.1. 从一付扑克牌中抽牌时, 可以用纸牌的花色和数字来说明其特征.

例 2.4.2. 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.

定义 2.4.1. 设$X=(X_1,\ldots,X_n)$. 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和连续型的.

定义 2.4.2. 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的所有可能取值为$\{a_{i1},a_{i2},\cdots\}$, $i=1,\ldots,n$, 则称

$$p(j_1,\cdots,j_n)=P(X_1=a_{1j_1},\ldots,X_n=a_{nj_n}),\ j_1,...,j_n=1,2,... \tag{2.4.1}$$

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

例 2.4.3. 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$M(N;p_1,\ldots,p_n)$.

解: 由于试验独立进行, 总的结果数为N，记结果A_i出现的次数为k_i，则$k_1+\cdots+k_n=N$。因此相当于多组组合，所以

$$\begin{aligned} P(X_1=k_1,\cdots,X_n=k_n) &= \frac{N!}{k_1!\cdots k_n!}P(\underline{A_1\cdots A_1}\cdots\underline{A_n\cdots A_n}) \\ &= \frac{N!}{k_1!\cdots k_n!}p_1^{k_1}\cdots p_n^{k_n}, \end{aligned}$$

其中 $k_1,\ldots,k_n$ 为非负整数且 $k_1+\cdots+k_n=N$.

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$\{(x_i,y_j):i=1,...,n,j=1,2,...,m\}$. 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)的概率函数可以下表表示:

$Y \backslash X$	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

例 2.4.4. *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*

$$p(x,y)=\frac{\binom{6}{x}\binom{7}{y}\binom{5}{4-x-y}}{\binom{18}{4}},\ 0\le x+y\le 4. \tag{2.4.2}$$

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.

定义 2.4.3. 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

$$P(a_1 \le X_1 \le b_1, ..., a_n \le X_n \le b_n) = \int_{a_n}^{b_n} ... \int_{a_1}^{b_1} f(x_1, \ldots, x_n) dx_1 \cdots dx_n, \tag{2.4.3}$$

则称f为X的概率密度函数.

对n维随机变量我们也有分布函数的概念.

定义 2.4.4. 设$X = (X_1, \ldots, X_n)$为n维随机变量. 对任意的$(x_1, \ldots, x_n) \in \mathbb{R}^n$, 称

$$F(x_1, \ldots, x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n) \tag{2.4.4}$$

为n维随机变量X的(联合)分布函数.

可以验证分布函数$F(x_1, \ldots, x_n)$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,

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

对高维离散型随机变量, 一般我们不使用分布函数.

例 2.4.5. 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.

例 2.4.6. 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$-\infty<a,b<\infty,\ 0<\sigma_1,\sigma_2<\infty,\ -1\le\rho\le 1$. 称$(X,Y)$服从参数为$a,b,\sigma_1,\sigma_2,\rho$的二元正态分布，记为$N(a,b,\sigma_1^2,\sigma_2^2,\rho)$.

§2.5 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$X_{i_1},...,X_{i_m}$为$X_1,...,X_n$的任一子集，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.

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$\{(x_i,y_j): i,j=1,2,\cdots\}$，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 \ 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

$$p_X(x_i) = P(X = x_i) = \sum_{j}^{m} P(X = x_i, Y = y_j) = \sum_{j}^{m} p_{ij} = p_{i\cdot},\ i = 1, 2, \cdots n. \tag{2.5.1}$$

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(2.5.1)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. 所以从列联表中, 我们不仅得到两个随机变量的联合分布, 同时通过将每行和每列相加, 得到两个变量的边缘分布.

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$X_{i_1}, ..., X_{i_m}$为$X_1, ..., X_n$的任一子集，则$X_{i_1}, ..., X_{i_m}$的概率函数为

$$p_{i_1 \ldots i_m}(j_{i_1}, ..., j_{i_m}) = P(X_{i_1} = a_{i_1 j_{i_1}}, ..., X_{i_m} = a_{i_m j_{i_m}}) = \sum p(j_1, ..., j_n).$$

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

例 2.5.1. *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.*

解：简单计算得到

$\eta\backslash\xi$	0	1	$p_{\cdot j}$
0	9/25	6/25	3/5
1	6/25	4/25	2/5
$p_{i\cdot}$	3/5	2/5	1

$\eta\backslash\xi$	0	1	$p_{\cdot j}$
0	6/20	6/20	3/5
1	6/20	2/20	2/5
$p_{i\cdot}$	3/5	2/5	1

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

$$P(x_1 \le X \le x_2) \quad = \quad P(x_1 \le X \le x_2, -\infty < Y < +\infty)$$

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty}\int_{x_1}^{x_2} f(u,v)dudv \\
&= \int_{x_1}^{x_2} f_X(u)du, \qquad (2.5.2)
\end{aligned}
$$

其中

$$f_X(u) = \int_{-\infty}^{+\infty} f(u,v)dv. \qquad (2.5.3)$$

从(2.5.2)我们可以看出, X的边缘密度函数即为(2.5.3). 类似地, Y的边缘密度函数为

$$f_Y(u) = \int_{-\infty}^{+\infty} f(u,v)du. \qquad (2.5.4)$$

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, $X_{i_1},...,X_{i_m}$的概率密度函数为

$$f(x_{i_1},...,x_{i_m}) = \int ... \int f(x_1,...,x_n)dx_1...dx_n.$$

其中积分是对除$X_{i_1},...,X_{i_m}$之外的所有变量来求积.

例 2.5.2. 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$N(b,\sigma_2^2)$.

例2.5.2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.

例 2.5.3. 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。

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。**

例 2.5.4. 设(X, Y)的联合概率密度有形式$(\forall (x, y) \in \mathbb{R}^2)$

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

其中$-\infty < a, b < \infty; 0 < \sigma_1, \sigma_2 < \infty; -1 \leq \rho \leq 1$. 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:

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-a)^2}{\sigma_1^2} - 2\rho\frac{(x-a)(y-b)}{\sigma_1\sigma_2} + \frac{(y-b)^2}{\sigma_2^2}\right]\right\} dy \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} exp\left\{-\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)}\right\} dv \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1} exp\{-\frac{1}{2}[(\frac{v-\rho u}{\sqrt{1-\rho^2}})^2 + u^2]\} dv \\
&= \frac{1}{\sqrt{2\pi}\sigma_1} exp\{-\frac{(x-a)^2}{2\sigma_1^2}\}
\end{aligned}$$

即$X \sim N(a, \sigma_1^2)$. 类似可得$Y \sim N(b, \sigma_2^2)$, 其边际概率密度为$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} exp\{-\frac{(y-b)^2}{2\sigma_2^2}\}$.