第二章随机变量及其分布

2.4 条件分布和随机变量的独立性

2.4.1 条件分布

一个随机变量 (或向量) 的条件概率分布，就是在给定 (或已知) 某种条件 (某种信息) 下该随机变量 (向量) 的概率分布。

1. 离散型随机变量的条件分布

设 (X, Y) 为二维离散型随机变量，其全部的可能取值为 $\{(x_i, y_j) : i, j = 1, 2, \cdots\}$。记其联合分布律为

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

若对给定的事件 $\{Y = y_j\}$，其概率 $P(Y = y_j) > 0$，则称

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

为在给定 $Y=y_j$ 的条件下 X 的**条件分布律 (概率函数)**。类似的，若 $P(X=x_i)>0$，则称

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

为在给定条件 $X=x_i$ 下 Y 的**条件分布律**。

设二维随机向量 (X_1, X_2) 的联合分布律如下所示:

$X_1 \backslash X_2$	-1	0	5	行和 $p_{i\cdot}$
1	0.17	0.05	0.21	0.43
3	0.04	0.28	0.25	0.57
列和 $p_{\cdot j}$	0.21	0.33	0.46	1.00

试求当 $X_2 = 0$ 时，X_1 的条件分布律。

↑Example ↓Example

解:

设 $X=(X_1,X_2,\cdots,X_n)\sim M(N;p_1,p_2,\ldots,p_n)$，试求 X_1 在给定 $X_2=k$ 的条件下的条件分布律。

↑Example

↓Example

解：

2．连续型随机变量的条件分布

设 (X,Y) 有概率密度 $f(x,y)$，我们考虑在给定 $y \le Y \le y+\epsilon$ 的条件下 X 的条件分布函数 (设$P\{y \le Y \le y+\epsilon\} > 0$)

$$\begin{aligned} P(X \le x | y \le Y \le y+\epsilon) &= \frac{P(X \le x, y \le Y \le y+\epsilon)}{P(y \le Y \le y+\epsilon)} \\ &= \int_{-\infty}^{x} \int_{y}^{y+\epsilon} f(u,v)dvdu \Big/ \int_{y}^{y+\epsilon} f_Y(y)dy \\ &= \int_{-\infty}^{x} \frac{\int_y^{y+\epsilon} f(u,v)dv}{\int_y^{y+\epsilon} f_Y(y)dy} du \end{aligned}$$

对上式两端关于 x 求导并令 $\epsilon \to 0$, 可求得 X 在给定条件 $Y=y$ 下的条件概率密度为

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}, \qquad f_Y(y) > 0.$$

记为

$$X|Y=y \sim f_{X|Y}(x|y).$$

类似地有 Y 在给定 $X = x$ 的条件下的条件概率密度:

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}, \qquad f_X(x) > 0.$$

记为

$$Y|X = x \sim f_{Y|X}(y|x).$$

设 (X,Y) 服从二元正态分布 $N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，试求 $X|Y=y$ 的条件概率密度。

↑Example

↓Example

解：

设 X, Y 服从单位圆上的均匀分布，试求 $f_{X|Y}(x|y)$ 和 $f_{Y|X}(y|x)$.

↑Example

↓Example

解:

3. 更一般情形

无论离散型还是连续型条件分布，上述 (X, Y) 中的 X 和 Y 皆可推广到高维。例如：设 $(X_1, X_2, \cdots, X_n) \sim f(x_1, x_2, \ldots, x_n)$，且 $(X_1, \cdots, X_k) \sim g(x_1, \ldots, x_k)$，则可定义在 $(X_1, \cdots, X_k) = (x_1, \ldots, x_k)$ 的条件下，$(X_{k+1}, \cdots, X_n)$ 的条件密度为：

$$h(x_{k+1}, \ldots, x_n | x_1, \ldots, x_k) = \frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_k)}, \quad \text{其中 } g(x_1, \ldots, x_k) > 0.$$

注: 若记 $(X_1, \cdots, X_k) = \boldsymbol{X}$，$(X_{k+1}, \cdots, X_n) = \boldsymbol{Y}$，$(x_1, \ldots, x_k) = \boldsymbol{x}$，$(x_{k+1}, \ldots, x_n) = \boldsymbol{y}$，则上式还可表示为:

$$h(\boldsymbol{y}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}, \boldsymbol{y})}{g(\boldsymbol{x})}, \ g(\boldsymbol{x}) > 0$$

2.4.2 随机变量的独立性

直观地, 若条件分布等于无条件分布，或者说条件分布与“条件”无关，例如，设 $f_{X|Y}(x|y) = g(x)$，则可推出 $g(x) = f_1(x)$，从而得

到：

$$f(x,y) = f_1(x)f_2(y), \quad (x,y) \in \mathbb{R}^2$$

此时我们称 X 与 Y 是 (相互) 独立的。若随机变量 X, Y 相互独立，那么对任何区域 A, B，应该有

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

从而对任意 n 个随机变量 $X_1, \ldots, X_n$，可以通过如下方式定义独立性:

称随机变量 $X_1, \ldots, X_n$ 相互独立，如果对任意的实数区间 $A_1, \ldots, A_n$ 都有

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n)$$

Definition

在离散型场合:

称离散型随机变量 $X_1, \cdots, X_n$ 相互独立, 若它们的联合分布律等于各自的边缘分布律的乘积, 即

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n),$$

其中 $(x_1, \ldots, x_n)$ 为 $(X_1, X_2, \cdots, X_n)$ 的值域中的任意一点.

Definition

连续型场合:

称连续型随机变量 $X_1, \cdots, X_n$ 相互独立，若它们的联合密度等于各自的边缘密度的乘积，即

$$f(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n), \quad \forall \quad (x_1, \ldots, x_n) \in \mathbb{R}^n$$

Definition

一般的, 随机变量之间的独立性可以通过分布函数来刻画:

Definition

设 $X_1, \cdots, X_n$ 为 n 个随机变量，如果它们的联合分布函数等于各自边缘分布函数的乘积，即

$$F(x_1, \cdots, x_n) = F_1(x_1) \cdots F_n(x_n), \quad \forall\ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

则称随机变量 $X_1, \cdots, X_n$ 相互独立.

在离散型和连续型两种情况下, 可以证明本定义分别与定义2.4.2和定义2.4.2等价.

如果随机变量 $X_1,\cdots,X_n$ 相互独立，则容易证明其中任何一部分随机变量也相互独立. 然而一般来说, 仅由某一部分独立却无法推出 $X_1,\cdots,X_n$ 相互独立. 如见下例:

↑Example ↓Example

若 ξ,η 相互独立, 都服从 -1 和 1 这两点上的等可能分布，而 $\zeta=\xi\eta$。则 ζ,ξ,η 两两独立但不相互独立。

↑Example ↓Example

设 $(X,Y) \sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$，则 X 与 Y 相互独立的充要条件是 $\rho = 0$。

↑Example ↓Example

设 (X,Y) 服从矩形 $D = [a,b] \times [c,d]$ 上的均匀分布，则 X 与 Y 相互独立。

↑Example ↓Example

设 (X,Y) 服从单位圆上的均匀分布，则 X 与 Y 不独立。

↑Example ↓Example

设有 n 个事件：A_1, A_2, $\cdots$, A_n，对于每个事件 A_i，定义：$X_i = I_{A_i}$ (A_i 的示性函数), $i = 1, 2, \cdots, n$，则可证明：A_1 ,A_2, $\cdots$, A_n 独立 $\Longleftrightarrow$ $X_1, X_2, \cdots, X_n$ 独立。