Lecture 08: Mult-dimensional Random Variable

Distribution Function of 2-Dimensional Random Variable

$\boxed{\text{Definition 8.1}}$ 设 X, Y 是定义在 $(\Omega, \mathscr{F}, P)$ 上的随机变量，则称 (X, Y) 为 $(\Omega, \mathscr{F}, P)$ 上的**二维随机变量**。对于任意 $x, y \in \mathbb{R}$，称

$$F(x, y) = P(\{X \le x\} \cap \{Y \le y\}) = P(X \le x, Y \le y)$$

为 (X, Y) 的 **(联合) 分布函数**。

注：从图像上来看，$F(x_0, y_0)$ 即为落在点 (x_0, y_0) 左下方无穷矩形的概率。

$F(x, y)$ 的若干性质：

- 固定 x_0，$F(x_0, y)$ 单调不减；固定 y_0，$F(x, y_0)$ 单调不减。

  (推论：$\forall x_1 > x_2, y_1 > y_2, F(x_1, y_1) \ge F(x_2, y_2)$。)
- 固定 x_0，$\lim_{y \to -\infty} F(x_0, y) = F(x_0, -\infty) = 0$，类似可得 $F(-\infty, -\infty) = 0, F(+\infty, +\infty) = 1$。
- 固定 x_0，$F(x_0, y)$ 处处有左极限且右连续 (càdlàg)，固定 y_0 时亦然。
- $\forall x_1 > x_2, y_1 > y_2, F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1) + F(x_2, y_2) \ge 0$。

注：上述性质也是一个函数是某个 (X, Y) 的分布函数的充分条件。

$\boxed{\text{Definition 8.2}}$ 给定 (X, Y)，X 或 Y 的分布函数称为**边缘概率**。

在 $F(x, y)$ 已知的情况下，边缘概率可以直接求出：

$$F_X(x) = P(X \le x) = P(X \le x, Y \in \mathbb{R}) = \lim_{y \to +\infty} F(x, y)$$

$\boxed{\text{Definition 8.3}}$ 设随机变量 X, Y 满足对于任意 $x, y \in \mathbb{R}$，$\{X \le x\}$ 与 $\{Y \le y\}$ 独立，即 $P(X \le x, Y \le y) = P(X \le x)P(Y \le y), F(x, y) = F_x(x)F_Y(y)$，则称 X, Y 相互独立。

注：虽然上述定义只对一部分事件的独立性进行了描述，但更复杂的事件的情形也是可以推出的：

- $$\begin{aligned} P(x_1 < X \le x_2, y_1 < Y \le y_2) &= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \\ &= F_X(x_2)F_Y(y_2) - F_X(x_2)F_Y(y_1) - F_X(x_1)F_Y(y_2) + F_x(x_1)F_Y(y_1) \\ &= (F_X(x_2) - F_X(x_1))(F_Y(y_2) - F_Y(y_1)) \\ &= P(x_1 < X \le x_2)P(y_1 < Y \le Y_2) \end{aligned}$$
- 设 I_n, J_m 是两列左开右闭区间，且 $\forall n \le k, I_n \cap I_k = \emptyset, J_n \cap J_k = \emptyset$，那么

$$
\begin{aligned}
P(X \in \bigcup_{n=1}^{\infty} I_n, Y \in \bigcup_{m=1}^{\infty} J_m) &= P(\bigcup_{n=1}^{\infty}\bigcup_{m=1}^{\infty}\{X \in I_n, Y \in J_m\}) \\
&\overset{\text{可列可加性}}{=} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} P(X \in I_n, Y \in J_m) \\
&\overset{\text{上一条结论}}{=} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} P(X \in I_n)P(Y \in J_m) \\
&= \left(\sum_{n=1}^{\infty} P(X \in I_n)\right)\left(\sum_{m=1}^{\infty} P(Y \in J_m)\right) \\
&= P(X \in \bigcup_{n=1}^{\infty} I_n)P(Y \in \bigcup_{m=1}^{\infty} J_m)
\end{aligned}
$$

$\boxed{\text{Theorem 8.4}}$ 若 X, Y 独立，$f(x), g(y)$ 为连续函数或分段连续函数，那么 $f(X)$ 和 $g(Y)$ 也相互独立。

n-Dimensional Case

高维情况和二维情况的定义，性质基本相同。求解其 k 维边缘概率时，将剩下的 $n-k$ 维都推到 $+\infty$ 即可。

n 维随机变量的独立性要求：
$\forall(x_1, x_2, \cdots x_n) \in \mathbb{R}^n, P(X_1 \le x_1, \cdots X_n \le x_n) = \prod_{k=1}^{n} P(X_k \le x_k)$。联想 n 个事件独立的定义，该条件似乎更弱：n 个事件的独立性对所有子集都做了约束。但事实上上述条件也对所有子集做了约束：只要将某些维的 x_i 推到 $+\infty$，则是一个对子集的约束。

2-Dimensional Discrete Random Variable

$\boxed{\text{Definition 8.5}}$ 若二维随机变量 (X, Y) 的可能取值是有限多个或可列无限个，则称 (X, Y) 为离散型二维随机变量，设 (X, Y) 所有可能取值为 $(x_i, y_j), \forall i, j = 1, 2, \cdots$，则称
$P(X = x_i, Y = y_j), \forall i, j = 1, 2, \cdots$ 为 (X, Y) 的联合分布律，可用表格表示为

X/Y	y_1	$\cdots$	y_k	$\cdots$
x_1	p_{11}	$\cdots$	p_{1k}	$\cdots$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\cdots$
x_k	p_{k1}	$\cdots$	p_{kk}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

(X, Y) 联合分布律的性质：

- $\forall i, j, p_{ij} \ge 0$
- $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1$
- 边缘分布的求法：

$$
\begin{aligned}
P(X = x_i) = P\left(X = x_i, Y \in \bigcup_{j=1}^{\infty}\{y_j\}\right) &= P\left(\bigcup_{j=1}^{\infty}\{X = x_i, Y = y_j\}\right) \\
= \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) &= \sum_{j=1}^{\infty} p_{i,j} \overset{def}{=} P_{i\cdot} \\
P(Y = y_j) = \sum_{i=1}^{\infty} P(X = x_i, Y = y_j) &\overset{def}{=} P_{\cdot j}
\end{aligned}
$$

再考虑离散型二维随机变量的分布函数及其边缘分布函数：

$$F(x,y)=\sum_{i:x_i\le x}\sum_{j:y_j\le y}p_{i,j}$$

$$F_X(x)=\sum_{i:x_i\le x}p_{i\cdot}=\sum_{i:x_i\le x}\sum_{j=1}^{\infty}p_{i,j}$$

一般地，对于区域 $D\subset\mathbb{R}^2$，有

$$P((x,y)\in D)=\sum_{i,j:(x_i,y_j)\in D}p_{i,j}$$

对于离散型随机变量的独立性，X 和 Y 独立当且仅当对于任意 i,j，$P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$，或者写成 $p_{i,j}=p_{i\cdot}p_{\cdot j}$。