

第三章 多元随机变量及其分布

二元离散型随机变量

(一) 联合概率分布

定义: 若二元随机变量 (X, Y) 全部可能取到的不同值是有限对或可列有限对, 则称 (X, Y) 是离散型随机变量.

$$P(X=x_i, Y=y_j) = P_{ij}, \quad i, j = 1, 2, \dots$$

$$1^\circ P_{ij} \geq 0$$

$$2^\circ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_{ij} = 1$$

[Ex] 设随机变量 X 在 1, 2, 3, 4 四个整数中等可能地取一值, 另一个随机变量 Y 在 $1 \sim X$ 中等可能地取一整数, 试求 (X, Y) 的联合概率分布.

$$P(X=i, Y=j) = P(X=i)P(Y=j|X=i) = \frac{1}{4} \times \frac{1}{i} = \frac{1}{4i}$$

则 (X, Y) 的联合概率分布为:

$X \backslash Y$	1	2	3	4	$P(X=x_i)$
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$P(Y=y_j)$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	

(二) 边际分布

X, Y 的边际分布律为

$$P(Y=y_j) = P(X < +\infty, Y=y_j) = \sum_{i=1}^{\infty} P_{ij} \stackrel{\text{def}}{=} P_{.j}, \quad j=1, 2, \dots$$

$$P(X=x_i) = P(X=x_i, Y < +\infty) = \sum_{j=1}^{\infty} P_{ij} \stackrel{\text{def}}{=} P_{i.}, \quad i=1, 2, \dots$$

(三) 条件分布

若 $P(Y=y_j) > 0$, 则称

$$P(X=x_i | Y=y_j) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)} = \frac{P_{ij}}{P_{.j}}, \quad i=1, 2, \dots$$

为 $Y=y_j$ 条件下, 随机变量 X 的条件分布律

[Ex] (X, Y) 的联合分布律为

$X \backslash Y$	-1	0	1
1	0	0.2	0.2
2	0.1	0.1	0.4

求 ① $\{X=2\}$ 条件下 Y 的条件分布律

② $\{X+Y=2\}$ 条件下 X 的条件分布律

$$① P(X=2) = 0.6$$

$$P(Y=j | X=2) = \begin{cases} \frac{1}{6}, & j=-1 \\ \frac{1}{6}, & j=0 \\ \frac{2}{3}, & j=1 \end{cases}$$

$$② P(X+Y=2) = 0.3$$

$$P(X=i | X+Y=2) = \frac{P(X=i, Y=2-i)}{P(X+Y=2)} = \begin{cases} \frac{2}{3}, & i=1 \\ \frac{1}{3}, & i=2 \end{cases}$$

[Ex] 求分布函数

$X \backslash Y$	0	1
0	0.1	0.2
1	0.3	0.4

$$F(x, y) = \begin{cases} 0, & x < 0 \text{ or } y < 0 \\ 0.1, & 0 \leq x < 1 \text{ or } 0 \leq y < 1 \\ 0.3, & 0 \leq x < 1 \text{ or } y \geq 1 \\ 0.4, & 0 \leq y < 1 \text{ or } x \geq 1 \\ 1, & x \geq 1, y \geq 1 \end{cases}$$

二元随机变量的分布函数

(一) 分布函数

定义: 设 (X, Y) 是二元随机变量, 对于任意实数 x, y ,

二元随机变量 (X, Y) 的分布函数为

$$F(x, y) = P\{(X \leq x) \cap (Y \leq y)\} \stackrel{\text{def}}{=} P(X \leq x, Y \leq y)$$

分布函数 $F(x, y)$ 的性质:

1° $F(x, y)$ 关于 x, y 单调不减, 即:

$$x_1 < x_2 \Rightarrow F(x_1, y) \leq F(x_2, y)$$

$$y_1 < y_2 \Rightarrow F(x, y_1) \leq F(x, y_2)$$

$$2^\circ 0 \leq F(x, y) \leq 1, F(+\infty, +\infty) = 1$$

对任意 x, y

$$F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0$$

3° $F(x, y)$ 关于 x, y 右连续, 即:

$$\lim_{\varepsilon \rightarrow 0^+} F(x+\varepsilon, y) = F(x, y)$$

$$\lim_{\varepsilon \rightarrow 0^+} F(x, y+\varepsilon) = F(x, y)$$

4° 若 $x_1 < x_2, y_1 < y_2$

$$\Rightarrow F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$$

$$\text{因为 } P(x_1 < X \leq x_2, y_1 < Y \leq y_2) =$$

$$F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$$

(二) 边际分布函数

$$\begin{cases} F_X(x) = F(x, +\infty) \\ F_Y(y) = F(+\infty, y) \end{cases}$$

(三) 条件分布函数

若 $P(Y=y) > 0$, 则 $Y=y$, X 的条件分布函数为

$$F_{X|Y}(x|y) = P(X \leq x | Y=y) = \frac{P(X \leq x, Y=y)}{P(Y=y)}$$

二元连续型随机变量

(一) 联合概率密度

定义: 对于二元随机变量 (X, Y) 的分布函数 $F(x, y)$,

如果存在非负函数 $f(x, y)$, 使对于任意 x, y ,

$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv$$

称 (X, Y) 为二元连续型随机变量

称 $f(x, y)$ 为二元随机变量 (X, Y) 的

(联合) 概率密度。

概率密度的性质:

1. $f(x, y) \geq 0$,

2. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$,

3. 设 G 是 xoy 平面上的区域, 点 (X, Y) 落在 G 内的概率为:

$$P((X, Y) \in G) = \iint_G f(x, y) dx dy$$

4. 在 $f(x, y)$ 的连续点 (x, y) , 有 $\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$.

【例1】设二元随机变量 (X, Y) 具有概率密度:

$$f(x, y) = \begin{cases} ke^{-(x+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

(1) 求常数 k

(2) 求分布函数 $F(x, y)$

(3) 求 $P(Y \leq X)$ 的概率.

(1) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$

$$\Leftrightarrow k \int_0^{+\infty} \int_0^{+\infty} e^{-x-y} dx dy = \frac{1}{6} \Rightarrow k = 6$$

(2) $F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv$

$$= \begin{cases} (1-e^{-x})(1-e^{-y}), & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

(3) $P(Y \leq X) = \int_0^{+\infty} \int_y^{+\infty} 6e^{-(x+y)} dx dy = \frac{3}{5}$

【例2】设二元随机变量 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} cy, & 0 < x < 1, x^2 < y < x \\ 0, & \text{else} \end{cases}$$

求 (1) 常数 C ; (2) $P(X > 0.5)$;

(3) $P(Y \leq 0.5)$; (4) $P(X > 0.5, Y \leq 0.5)$

(1) $\int_0^1 dx \int_{x^2}^x cy dy = \frac{c}{15} = 1 \Rightarrow c = 15$

(2) $P(X > 0.5) = \int_{0.5}^1 dx \int_{x^2}^x 15y dy = \frac{47}{84} = 0.734$

(3) $P(Y \leq 0.5) = \int_0^{\frac{1}{2}} dy \int_y^{+\infty} 15y dx = \frac{6\sqrt{2}-5}{8} = 0.436$

(4) $P(X > 0.5, Y \leq 0.5)$

$$= \int_{\frac{1}{4}}^{\frac{1}{2}} dy \int_{0.5}^y 15y dx = 0.170$$

(5) 求 X, Y 的边缘概率密度 $f_X(x), f_Y(y)$

$$f_X(x) = \begin{cases} \int_{x^2}^x 15y dy, & 0 < x < 1 \\ 0, & \text{else} \end{cases}$$

$$= \begin{cases} \frac{15}{2}(x^2 - x^4), & 0 < x < 1 \\ 0, & \text{else} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_y^{+\infty} 15y dx = 15(y^{\frac{2}{3}} - y^2), & 0 < y < 1 \\ 0, & \text{else} \end{cases}$$

(二) 边际概率密度

(二) 边际 (边缘) 概率密度

对于连续型随机变量 (X, Y) , 概率密度为 $f(x, y)$,

X, Y 的边际概率密度为:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

二元均匀分布与二元正态分布

若二元随机变量 (X, Y) 在二维有界区域 D 上取值, 且具有概率密度

$$f(x, y) = \begin{cases} \frac{1}{D \text{ 的面积}}, & (x, y) \in D \\ 0, & \text{其他} \end{cases}$$

则称 (X, Y) 在 D 上服从均匀分布 (Uniform Distribution)

二元正态分布 设二元随机变量 (X, Y) 的概率密度为:

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

$$(-\infty < x < +\infty, -\infty < y < +\infty)$$

其中 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 都是常数, 且 $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$

称 (X, Y) 为服从参数为 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 的二元正态分布,

记为: $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

随机变量的独立性

设 $F(x, y)$ 及 $F_X(x), F_Y(y)$ 分别是二元随机变量 (X, Y) 的分布函数及边际分布函数,

若对所有 x, y 有: $P(X \leq x, Y \leq y) = P(X \leq x) P(Y \leq y)$

即 $F(x, y) = F_X(x) F_Y(y)$, 则称随机变量 X, Y 相互独立

连续型随机变量 X, Y 相互独立, 其密度函数应有 $f(x, y) = g(x) h(y), a < x < b, c < y < d$.

二元随机变量的函数的分布

设二元离散型随机变量 (X, Y) 具有概率分布 $P(X=x_i, Y=y_j) = p_{ij}, i, j = 1, 2, \dots$

(1) 设 $U = u(X, Y), V = v(X, Y)$

(2) $Z = g(X, Y)$

对于 (1), 先确定 (U, V) 的取值 $(u_i, v_j), i, j = 1, 2, \dots$

再找出 $(U = u_i, V = v_j) = \{(X, Y) \in D\}$, 从而计算出分布律;

对于 (2) 类似 (1), 先确定 Z 的取值 $z_i, i = 1, 2, \dots$

再找出 $(Z = z_i) = \{(X, Y) \in D\}$, 从而计算出分布律;

卷积公式: 当 X 与 Y 相互独立时, $Z = X + Y$ 的密度函数公式

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx$$

[EX] 设二元随机变量 (X, Y) 在区域 $\{(x, y): |y| < x < 1\}$ 内均匀分布, 求条件概率密度 $f_{X|Y}(x|y)$ 及 $P(X > \frac{2}{3} | Y = \frac{1}{2})$

$$f(x, y) = \begin{cases} 1, & |y| < x < 1 \\ 0, & \text{else} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_{|y|}^1 dx = 1-|y|, & -1 < y < 1 \\ 0, & \text{else} \end{cases}$$

$$\Rightarrow f_{X|Y}(x|y) = \begin{cases} \frac{1}{1-|y|}, & |y| < x < 1 \\ 0, & \text{else} \end{cases}$$

$$P(X > \frac{2}{3} | Y = \frac{1}{2}) = \int_{\frac{2}{3}}^1 2dx = \frac{2}{3}$$

[EX] 设二元随机变量 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

求 (1) X, Y 的边缘概率密度 $f_X(x), f_Y(y)$.

(2) 条件概率密度 $f_{Y|X}(y|x), f_{X|Y}(x|y)$

$$(1) f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, -\infty < x < +\infty, \text{ 即 } X \sim N(\mu_1, \sigma_1^2)$$

对于 $f_Y(y)$ 同理

$$(2) f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \dots$$

[EX] (X, Y) 具有概率密度 $f(x, y) = \begin{cases} 6e^{-2x-3y}, & x > 0, y > 0 \\ 0, & \text{else} \end{cases}$

其中 X 与 Y 是否相互独立

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & \text{else} \end{cases} \quad f_Y(y) = \begin{cases} 3e^{-3y}, & y > 0 \\ 0, & \text{else} \end{cases}$$

故有 $f(x, y) = f_X(x) \cdot f_Y(y)$, 因而 X, Y 是相互独立的