

第四章、随机向量

§4.1 随机向量的联合分布与边缘分布

- 需要研究多个随机变量的情况, 如:
弹着点横坐标 X 与纵坐标 Y ;
炼钢厂每炉钢的硬度 X 、含碳量 Y 、含硫量 Z .
- 重点: 变量之间有**联系**.
- 定义0.1. 称 n 个随机变量 X_1, X_2, \dots, X_n 的整体
 $\xi = (X_1, X_2, \dots, X_n)$ 为 n 维随机向量.
- 注: “维数” 是分量的个数. 如, 弹着点坐标 (X, Y) 是二维平面中的随机点.
- 注: 着重讨论二维随机向量, 高维类似.

二维离散型随机向量

- 定义1.1. 若 $\xi = (X, Y)$ 只取有限个或可列个(二维实向量)值, 则称 ξ 为离散型随机向量.

- 注: 等价地, 两个分量 X, Y 都是离散型随机变量.

- 取值范围:

$$X: \{x_i : i = 1, 2, \dots\}; \quad Y: \{y_j : j = 1, 2, \dots\},$$
$$(X, Y): \{(x_i, y_j) : i = 1, 2, \dots, j = 1, 2, \dots\}.$$

- $\xi = (X, Y)$ 的(概率)分布:

$$P((X, Y) = (x_i, y_j)) = p_{ij}, \quad i = 1, 2, \dots, j = 1, 2, \dots \quad (1.1)$$

也称为 (X, Y) 的联合分布(列).

- 注: 允许出现有些组合 (x_i, y_j) 对应的概率 $p_{ij} = 0$.

- 二维概率分布列表:

$X \setminus Y$	y_1	y_2	\cdots	y_j	\cdots
x_1	p_{11}	p_{12}	\cdots	p_{1j}	\cdots
x_2	p_{21}	p_{22}	\cdots	p_{2j}	\cdots
\vdots	\vdots	\vdots		\vdots	
x_i	p_{i1}	p_{i2}	\cdots	p_{ij}	\cdots
\vdots	\vdots	\vdots		\vdots	

- 形如 $\{(X, Y) = (x_i, y_j)\}$ 的事件构成完备事件组, 因此:

$$(1) p_{ij} \geq 0, \quad \forall i = 1, 2, \dots, j = 1, 2, \dots$$

$$(2) \sum_i \sum_j p_{ij} = 1.$$

例1.1. 设 (X, Y) 仅取5 个不同点:

$(1, 1)$ $(1.2, 1)$ $(1.4, 1.5)$ $(1, 1.3)$ $(0.9, 1.2)$

且取每个点的概率相等(均为 $\frac{1}{5}$).

• 联合分布:

概率分布表:

$$\begin{aligned}P((X, Y) = (1, 1)) &= \frac{1}{5}, \\P((X, Y) = (1, 2.1)) &= \frac{1}{5}, \\P((X, Y) = (1.4, 1.5)) &= \frac{1}{5}, \\P((X, Y) = (1, 1.3)) &= \frac{1}{5}, \\P((X, Y) = (0.9, 1.2)) &= \frac{1}{5}.\end{aligned}$$

$X \setminus Y$	1	1.2	1.3	1.5
0.9	0	$\frac{1}{5}$	0	0
1	$\frac{1}{5}$	0	$\frac{1}{5}$	0
1.2	$\frac{1}{5}$	0	0	0
1.4	0	0	0	$\frac{1}{5}$

例1.2. 三项分布(参数: n 为正整数, $0 < p_1, p_2 < 1$, $p_1 + p_2 < 1$).
 设 (X, Y) 的联合分布为

$$\begin{aligned}
 P((X, Y) = (k_1, k_2)) &\triangleq p_{k_1, k_2} \\
 &= \frac{n!}{k_1! k_2! (n - k_1 - k_2)!} p_1^{k_1} p_2^{k_2} (1 - p_1 - p_2)^{n - k_1 - k_2}, \\
 k_1, k_2 &= 0, 1, \dots, n, \quad k_1 + k_2 \leq n.
 \end{aligned} \tag{1.3}$$

- 验证所有概率之和等于1:

$$\begin{aligned}
 &\sum_{k_1=0}^n \sum_{k_2=0}^{n-k_1} p_{k_1, k_2} \quad (\text{固定 } k_1, \text{ 令 } n - k_1 = m) \\
 &= \sum_{k_1=0}^n \frac{n!}{k_1! (n - k_1)!} p_1^{k_1} \sum_{k_2=0}^m \frac{m!}{k_2! (m - k_2)!} p_2^{k_2} (1 - p_1 - p_2)^{m - k_2} \\
 &= \sum_{k_1=0}^n \frac{n!}{k_1! (n - k_1)!} p_1^{k_1} \cdot \left(p_2 + (1 - p_1 - p_2) \right)^{n - k_1} = 1.
 \end{aligned}$$

例1.3 (三项分布的实例). 一大批粉笔, 白色(W)占比: $p_W = 60\%$; 黄色(Y)占比: $p_Y = 25\%$, 红色(R)占比: $p_R = 15\%$. 随机(顺序地)取出6支. 问: 其中恰有3支白色、1支黄色、2支红色的概率.

- 解: 用长度为6的W-Y-R字符串表示(顺序地)抽取结果.
- 可以认为各次抽取独立, 且抽到各颜色的概率不变: 如,

$$P(\text{WWWWWW}) = p_W p_W p_W p_W p_W p_W = (0.6)^6,$$

$$P(\text{WWWYRR}) = p_W p_W p_W p_Y p_R p_R = (0.6)^3 (0.25)^1 (0.15)^2.$$

- 恰有3个W, 1个Y, 2个R的字符串数目:

$$C_6^3 C_3^1 = \frac{6!}{3!1!2!} = 60.$$

- 每个字符串对应的概率均为 $(0.6)^3 (0.25)^1 (0.15)^2$. 故所求为

$$60 \cdot (0.6)^3 (0.25)^1 (0.15)^2 = 0.0729.$$

- 令 $X =$ “6支中白粉笔数目”, $Y =$ “6支中黄粉笔数目”.
- 则事件 “6支中恰有3支白色、1支黄色、2支红色” 即为

$$\{X = 3, Y = 1\} = \{(X, Y) = (3, 1)\}.$$

- 上面的推导,

$$P((X, Y) = (3, 1)) = \frac{6!}{3!1!2!} (0.6)^3 (0.25)^1 (0.15)^2.$$

- 一般地, 对于 $0 \leq k_1, k_2 \leq 6, 0 \leq k_2 \leq 6, k_1 + k_2 \leq 6,$

$$\begin{aligned} & P((X, Y) = (k_1, k_2)) \\ &= \frac{6!}{k_1!k_2!(6 - k_1 - k_2)!} (0.6)^{k_1} (0.25)^{k_2} (0.15)^{6 - k_1 - k_2}. \end{aligned}$$

- (X, Y) 服从参数为 $n = 6, p_1 = 0.6, p_2 = 0.25$ 的三项分布.

- (X, Y) 的分布称为联合分布.
 X 的分布称为 (X, Y) 的关于 X 的边缘分布; Y 的类似.
- 联合分布决定边缘分布: 设

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

- 则

$$\begin{aligned} P(X = x_i) &= \sum_j P(X = x_i, Y = y_j) \quad (\text{完全/有限可加性}) \\ &= \sum_j p_{ij}; \quad (X \text{ 的边缘分布.}) \end{aligned} \quad (1.4)$$

$$P(Y = y_j) = \sum_i p_{ij}. \quad (Y \text{ 的边缘分布.}) \quad (1.4')$$

例1.4. 在例1.1 的概率分布表中, 分别按行、列求和, 就得到了 X 、 Y 的边缘分布:

$X \setminus Y$	1	1.2	1.3	1.5	
0.9	0	$\frac{1}{5}$	0	0	$\frac{1}{5}$
1	$\frac{1}{5}$	0	$\frac{1}{5}$	0	$\frac{2}{5}$
1.2	$\frac{1}{5}$	0	0	0	$\frac{1}{5}$
1.4	0	0	0	$\frac{1}{5}$	$\frac{1}{5}$
	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	

例1.5. 设 (X, Y) 服从参数为 n, p_1, p_2 的三项分布:

$$P((X, Y) = (k_1, k_2)) \triangleq p_{k_1, k_2}$$

$$= \frac{n!}{k_1! k_2! (n - k_1 - k_2)!} p_1^{k_1} p_2^{k_2} (1 - p_1 - p_2)^{n - k_1 - k_2},$$

$$k_1, k_2 = 0, 1, \dots, n, \quad k_1 + k_2 \leq n.$$

- $X \sim B(n, p_1)$, (类似地, $Y \sim B(n, p_2)$):

$$P(X = k_1) = \sum_{k_2=0}^{n-k_1} p_{k_1, k_2}$$

$$= \frac{n!}{k_1! (n - k_1)!} p_1^{k_1} \sum_{k_2=0}^{n-k_1} \frac{(n - k_1)!}{k_2! (n - k_1 - k_2)!} p_2^{k_2} (1 - p_1 - p_2)^{n - k_1 - k_2}$$

$$= \frac{n!}{k_1! (n - k_1)!} p_1^{k_1} \cdot (p_2 + (1 - p_1 - p_2))^{n - k_1}, \quad k_1 = 0, 1, \dots, n.$$

- 注: 将白色(W)视为成功, 将黄色(Y)和红色(R)均视为失败.

二维连续型随机向量

- 定义1.2. 若存在非负函数 $p(x, y)$ 使得: 对任意 $a < b, c < d$, (记 $D = \{(x, y) : a < x < b, c < y < d\}$),

$$P((X, Y) \in D) = \iint_D p(x, y) dx dy, \quad (1.5)$$

则称 $\xi = (X, Y)$ 为连续型随机向量. 称 $p(x, y)$ 为 (X, Y) 的联合分布密度, 简称联合密度或分布密度.

- 注: 可改变 $p(x, y)$ 在一条线上的函数值.
- 注: $p(x, y)$ 连续或分段连续. 若在 (x_0, y_0) 连续, 则
$$p(x_0, y_0) = \lim_{\delta \rightarrow 0} \frac{1}{\delta^2} P(x_0 < X < x_0 + \delta, y_0 < Y < y_0 + \delta).$$
- 注: 联合密度不是概率, 其积分才是概率.
- 注: 对二维区域 D , (1.5) 仍成立. 积分为曲顶主体的体积. 特别地, 取 D 为全平面, 则积分为1.

例1.6. 设 (X, Y) 的联合密度如下:

$$p(x, y) = \begin{cases} Ce^{-(x+y)} & \text{若 } x \geq 0, y \geq 0, \\ 0 & \text{其它.} \end{cases}$$

(1) 求常数 C ; (2) 求 $P(0 < X < 1, 0 < Y < 1)$.

● 解: (1) $C = 1$:

$$\begin{aligned} 1 &= \int_0^{\infty} \int_0^{\infty} p(x, y) dx dy = C \int_0^{\infty} \int_0^{\infty} e^{-(x+y)} dx dy \\ &= C \int_0^{\infty} e^{-x} dx \int_0^{\infty} e^{-y} dy = C. \end{aligned}$$

● (2) 记 $D = \{(x, y) : 0 < x < 1, 0 < y < 1\}$, 则所求为

$$\begin{aligned} \iint_D p(x, y) dx dy &= \int_0^1 \int_0^1 e^{-(x+y)} dx dy \\ &= \int_0^1 e^{-x} dx \int_0^1 e^{-y} dy = (1 - e^{-1})^2. \end{aligned}$$

- 定理1.1. 若 $\xi = (X, Y)$ 是连续型, 联合密度记为 $p(x, y)$. 则 X, Y 都是连续型, 且

$$p_X(x) = \int_{-\infty}^{\infty} p(x, y)dy, \quad p_Y(y) = \int_{-\infty}^{\infty} p(x, y)dx. \quad (1.7)$$

- 证: $\forall a < b$. 记 $D = \{(x, y) : a < x < b, -\infty < y < \infty\}$. 则

$$\begin{aligned} P(a < X < b) &= P(\xi \in D) \\ &= \iint_D p(x, y)dx dy = \int_a^b \int_{-\infty}^{\infty} p(x, y)dy dx. \end{aligned}$$

- 定义1.3. 设 $\xi = (X, Y)$ 为连续型. 称 $p_X(x)$ 为 ξ 关于 X 的边缘(分布)密度. 关于 Y 的类似.

- 定义1.4 设 G 是平面上的区域, 面积 $a \in (0, \infty)$. 若对 G 的任意子区域 A , $P((X, Y) \in A) = \frac{A \text{ 的面积}}{G \text{ 的面积}}$. 则称 (X, Y) 服从 G 上的均匀分布. 记为 $U(G)$.
- 注: (X, Y) 为连续型, 联合密度为

$$p(x, y) = \begin{cases} \frac{1}{a}, & \text{若 } (x, y) \in G, \\ 0, & \text{若 } (x, y) \notin G. \end{cases}$$

例1.7. 设 G 为平面上由 $y = x^2$ 和 $y = x$ 所夹的有限区域,
 $(X, Y) \sim U(G)$. 求 (X, Y) 的联合密度和边缘密度.

• 解: G 的面积为 $a = \int_0^1 x dx - \int_0^1 x^2 dx = \frac{1}{6}$.

• 联合密度为 $p(x, y) = \begin{cases} 6, & \text{若}(x, y) \in G, \\ 0, & \text{若}(x, y) \notin G. \end{cases}$

• 边缘密度: $p_X(x) = 0, (x \notin [0, 1]), p_Y(y) = 0, (y \notin [0, 1]);$

$$p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy = \int_{x^2}^x 6 dy = 6(x - x^2), \quad (x \in [0, 1]),$$

$$p_Y(y) = \int_{-\infty}^{\infty} p(x, y) dx = \int_y^{\sqrt{y}} 6 dx = 6(\sqrt{y} - y), \quad (y \in [0, 1]).$$

• 注: X, Y 均可以取遍 $[0, 1]$, 但 (X, Y) 不能取遍 $[0, 1] \times [0, 1]$.

随机变量的独立性

- 定义1.5. 若 $\forall a < b$ 和 $c < d$, $\{a < X < b\}$ 与 $\{c < Y < d\}$ 相互独立, 则称 X 与 Y 是相互独立的, 简称独立.

- 定理1.2(连续型). 设 X, Y 均为连续型. 则 X 与 Y 相互独立的充要条件是: $p_X(x)p_Y(y)$ 是随机向量 (X, Y) 的联合密度.

- 证: 令 $D = \{(x, y) : a < x < b, c < y < d\}$. 则

$$\begin{aligned} \iint_D p_X(x)p_Y(y)dx dy &= \int_a^b p_X(x)dx \cdot \int_c^d p_Y(y)dy = P(\star) \cdot P(\star) \\ &\quad \parallel ? \qquad \qquad \qquad \qquad \qquad \qquad \qquad \parallel ? \\ P((X, Y) \in D) &= P(\star, \star) \end{aligned}$$

- “ \Leftarrow ” : 上与下的左边相等, 故右边相等.
- “ \Rightarrow ” : 上与下的右边相等, 故左边相等.

- 定理1.3(离散型). 设 X 的所有可能取值为 x_1, x_2, \dots (有限或可列), Y 的为 y_1, y_2, \dots . 则 X 与 Y 相互独立的充要条件是:

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j), \quad \forall i, j. \quad (1.9)$$

- 证: 令 $D = \{(x, y) : a < x < b, c < y < d\}$,

$$A = \{x_i : a < x_i < b\}, \quad B = \{y_j : c < y_j < d\}.$$

- “ \Leftarrow ” : 上与下的左边相等, 故右边相等.

$$\sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i)P(Y = y_j) = P(\star) \cdot P(\star)$$

|| ? || ?

$$\sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i, Y = y_j) = P(\star, \star)$$

- “ \Rightarrow ” : 上与下的右边相等. 取 $a = x_i - \frac{1}{n}$, $b = x_i + \frac{1}{n}$,
 $c = y_j - \frac{1}{n}$, $d = y_j + \frac{1}{n}$. 然后令 $n \rightarrow \infty$.

例1.8. 设 $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, 且 X_1 与 X_2 相互独立, 求 (X_1, X_2) 的联合密度.

- 解: X_1, X_2 的密度分别为

$$X_1 \sim p_1(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{1}{2} \frac{(x_1 - \mu_1)^2}{\sigma_1^2}\right\},$$
$$X_2 \sim p_2(x_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{1}{2} \frac{(x_2 - \mu_2)^2}{\sigma_2^2}\right\}.$$

- 相互独立, 故联合密度为

$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2} \left[\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2 \right]\right\}.$$

- 注: 联合密度决定边缘密度, 反之不然.
若分量相互独立, 则反之亦然.

二维正态分布

- 定义1.6. 若 $\xi = (X, Y)$ 的联合密度如下, 则称 ξ 服从二维正态分布, 称 ξ 为二维正态随机向量.

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \cdot I \right\},$$

$$\text{其中, } I = \frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}$$
$$= u^2 - 2\rho uv + v^2. \quad u = \frac{x-\mu_1}{\sigma_1}, \quad v = \frac{y-\mu_2}{\sigma_2}.$$

- 其中 $\mu_1, \mu_2, \sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$ 是5个参数.
- 称 $p(x, y)$ 为二维正态密度.
- 注: “二维” 也可换成 “二元” .

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

- 联合密度: $p(x, y) = C \exp \left\{ -\frac{1}{2(1-\rho^2)} \cdot I \right\}$,

$$I = u^2 - 2\rho uv + v^2, \quad u = \frac{x - \mu_1}{\sigma_1}, \quad v = \frac{y - \mu_2}{\sigma_2}, \quad C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}.$$

- $p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy$. 改写 $I = (v - \rho u)^2 + (1 - \rho^2)u^2$. 故

$$\begin{aligned} p_X(x) &= C \int_{-\infty}^{\infty} \exp \left\{ -\frac{(v - \rho u)^2 + (1 - \rho^2)u^2}{2(1 - \rho^2)} \right\} dy \\ &= C\sigma_2 \cdot e^{-\frac{u^2}{2}} \cdot \int_{-\infty}^{\infty} \exp \left\{ -\frac{(v - \rho u)^2}{2(1 - \rho^2)} \right\} dv \\ &= C\sigma_2 \cdot e^{-\frac{u^2}{2}} \cdot \sqrt{2\pi(1 - \rho^2)} = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}. \end{aligned}$$

- 注: 进一步, 推出 $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1$.

分量独立的充要条件: $\rho = 0$.

- 联合密度: $p(x, y) = C \exp \left\{ -\frac{1}{2(1-\rho^2)} \cdot I \right\},$

$$I = u^2 - 2\rho uv + v^2, \quad u = \frac{x - \mu_1}{\sigma_1}, \quad v = \frac{y - \mu_2}{\sigma_2}, \quad C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}.$$

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

- “ \Leftarrow ” : 设 $\rho = 0$. 则交叉项 $2\rho uv$ 消失, 且 $1 - \rho^2 = 1$. 易见 $p(x, y) = p_X(x)p_Y(y)$.
- “ \Rightarrow ” : 若 X, Y 独立, 则如下定义的 $\hat{p}(x, y)$ 是联合密度.

$$\hat{p}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left\{ -\frac{1}{2}(u^2 + v^2) \right\}. \quad (1.11)$$

$p(x, y)$ 与 $\hat{p}(x, y)$ 都是连续的, 故完全相等. 特别地,

$$p(\mu_1, \mu_2) = p_X(\mu_1)p_Y(\mu_2) \Rightarrow C = \frac{1}{2\pi\sigma_1\sigma_2} \Rightarrow \rho = 0.$$

- 定义1.7. 称二元函数

$$F(x, y) = P(X \leq x, Y \leq y)$$

为 $\xi = (X, Y)$ 的(联合)分布函数. 记为 $F_{\xi}(x, y)$ 或 $F_{X, Y}(x, y)$.

- 注: 设 (X, Y) 为连续型. 若联合密度 $p(x, y)$ 在 (x_0, y_0) 连续, 则

$$\frac{\partial^2 F_{X, Y}}{\partial x \partial y}(x_0, y_0) = p(x_0, y_0).$$

- 注: 若 $\frac{\partial^2 F_{X, Y}(x, y)}{\partial x \partial y}$ 存在且连续, 则它是 (X, Y) 的联合密度, (故, (X, Y) 为连续型).

§4.2 两个随机变量的函数的分布

- 记 $Z = X + Y$. 用分布函数法: 令 $D = \{(x, y) : y \leq z - x\}$,

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = P((X, Y) \in D),$$

- 设 (X, Y) 为连续型, 密度为 $p(x, y)$. 则

$$P(Z \leq z) = \iint_D p(x, y) dx dy = \iint_{y \leq z-x} p(x, y) dx dy \quad (2.1)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} p(x, y) dy dx \quad (\text{二重积分化为累次积分})$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^z p(x, u-x) du dx \quad (\text{固定 } x. \text{ 令 } u = y + x)$$

$$= \int_{-\infty}^z \int_{-\infty}^{\infty} p(x, u-x) dx du. \quad (\text{积分交换次序})$$

- 设 (X, Y) 为连续型, 密度为 $p(x, y)$, 令 $Z = X + Y$. 则

$$F_Z(z) = P(Z \leq z) = \int_{-\infty}^z \int_{-\infty}^{\infty} p(x, u - x) dx du.$$

- 因此, Z 为连续型, 且密度为

$$p_Z(z) = \int_{-\infty}^{\infty} p(x, z - x) dx. \quad (2.2)$$

- 特别地, 若 X, Y 相互独立, 则

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(x)p_Y(z - x) dx = \int_{-\infty}^{\infty} p_X(z - y)p_Y(y) dy.$$

例2.1. 设 X 与 Y 相互独立, 都服从 $N(\mu, \sigma^2)$, 求 $X + Y$ 的密度.

- 解: 将 (X, Y) 的联合密度记为 $p(x, y)$. 则

$$p(x, z - x) = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{(x - \mu)^2 + (z - x - \mu)^2}{2\sigma^2} \right\}.$$

- 令 $t = x - \mu$. 指数中的分子:

$$\begin{aligned} \star &= t^2 + (z - 2\mu - t)^2 = 2t^2 - 2(z - 2\mu)t + (z - 2\mu)^2 \\ &= 2\left(t - \frac{z - 2\mu}{2}\right)^2 + \frac{(z - 2\mu)^2}{2}. \quad (\text{记 } \mu_0 = \frac{z - 2\mu}{2}) \end{aligned}$$

- $Z \sim N(2\mu, 2\sigma^2)$: $p_Z(z) = \int_{-\infty}^{\infty} p(x, z - x)dx$, 故

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2} \cdot \exp \left\{ -\frac{(t - \mu_0)^2}{2 \cdot \sigma^2/2} \right\} \cdot \exp \left\{ -\frac{(z - 2\mu)^2}{2 \cdot 2\sigma^2} \right\} dt \\ &= \frac{1}{2\pi\sigma^2} \cdot \sqrt{2\pi\sigma/\sqrt{2}} \cdot e^{-\frac{(z-2\mu)^2}{2 \cdot 2\sigma^2}} = \frac{1}{\sqrt{2\pi}2\sigma^2} e^{-\frac{(z-2\mu)^2}{2 \cdot 2\sigma^2}}. \end{aligned}$$

分布函数法.

- 设 (X, Y) 联合密度为 $p(x, y)$, $Z = f(X, Y)$. 求 Z 的分布.
- 第一步、求 Z 的分布函数. $D_z = \{(x, y) : f(x, y) \leq z\}$.

$$F_Z(z) = P(f(X, Y) \leq z) = \iint_{D_z} p(x, y) dx dy.$$

- 第二步、直接计算 $F_Z(z)$, 然后求导得到 $p_Z(z)$.
或者, 利用变量替换、累次积分、积分交换次序等工具, 将积分化为最终形式.

$$F_Z(z) = \iint_{f(x, y) \leq z} p(x, y) dx dy = \cdots = \int_{-\infty}^z p_Z(u) du.$$

例2.2. 设 X, Y 独立, 都服从 $N(0, 1)$. 求 $Z = \sqrt{X^2 + Y^2}$ 的密度.

- 解: 若 $z \leq 0$, 则 $F_Z(z) = 0$. 若 $z > 0$, 则

$$F_Z(z) = \iint_{D_z} p(x, y) dx dy. \quad \text{其中,}$$

$$p(x, y) = \frac{1}{2\pi} \exp \left\{ -\frac{1}{2}(x^2 + y^2) \right\}, \quad D_z = \{(x, y) : x^2 + y^2 \leq z^2\}.$$

- 故, 作极坐标变换; 并计算雅可比行列式.

$$\begin{aligned} x &= r \cos \theta, & \begin{pmatrix} r \geq 0 \\ 0 \leq \theta < 2\pi \end{pmatrix}; & \frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r. \\ y &= r \sin \theta, \end{aligned}$$

$$\Rightarrow F_Z(z) = \int_0^{2\pi} \int_0^z \frac{1}{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta = \int_0^z r e^{-\frac{1}{2}r^2} dr, \quad (z > 0).$$

- 称 Z 服从瑞利(Rayleigh)分布, 密度如下:

$$p_Z(z) = z e^{-\frac{z^2}{2}}, \quad (z > 0); \quad p_Z(z) = 0, \quad (z \leq 0).$$

例2.3. 设 X, Y 独立同分布, 共同的密度函数为 $p(\cdot)$, 分布函数为 $F(\cdot)$. 求 $Z = \max(X, Y)$ 的密度函数.

● 解:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(\max(X, Y) \leq z) \\ &= P(X \leq z, Y \leq z) \\ &= P(X \leq z) \cdot P(Y \leq z) \quad (\text{独立性}) \\ &= F(z) \cdot F(z) = F^2(z). \end{aligned}$$

● 求导得到 Z 的密度:

$$p_Z(z) = (F^2(z))' = 2F(z) \cdot F'(z) = 2F(z)p(z).$$

二维变换后的联合密度

- 已知 (X, Y) 的联合密度. 求如下变换后的 (U, V) 的联合密度.

$$U = f(X, Y), \quad V = g(X, Y).$$

- 定理2.1. \star . 设区域 A 使得 $P((X, Y) \in A) = 1$; f, g 满足:

(1) 任意 (u, v) 在 A 中的原像(下述方程组的解)唯一.

$$f(x, y) = u, \quad g(x, y) = v. \quad (2.4)$$

(2) f, g 在 A 中有连续偏导数.

(3) 雅可比行列式 $J(x, y) = \frac{\partial(f, g)}{\partial(x, y)}$ 在 A 中处处不等于0.

设 \star . 令 $G = \{(u, v) : \text{方程组(2.4)在} A \text{ 中有(唯一)解}\}$. 则

$$q(u, v) = p(x, y) \cdot \frac{1}{|J(x, y)|}, \quad ((u, v) \in G); \quad q(u, v) = 0, \quad ((u, v) \notin G).$$

是 (U, V) 的联合密度, 其中 (x, y) 是(2.4) 在 A 中的(唯一)解.

定理2.1. 的证明:

- 固定 $a < b, c < d$. 记

$$D = \{(u, v) : a < u < b, c < v < d\},$$

$$D^* = \{(x, y) : (f(x, y), g(x, y)) \in D\}.$$

- (f, g) 是 $D^* \cap A$ 到 $D \cap G$ 上的一、一映射,

且 $\frac{\partial(x, y)}{\partial(u, v)} = 1 / \frac{\partial(u, v)}{\partial(x, y)}$. 由其逆映射的重积分的变数替换公式,

$$\iint_{D^* \cap A} p(x, y) dx dy = \iint_{D \cap G} p(x, y) \cdot \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv = \iint_{D \cap G} q(u, v) du dv.$$

其中, $q(u, v) = p(x, y) \cdot \frac{1}{|J(x, y)|}, \forall (u, v) \in G$.

- $q(u, v)$ 是 (U, V) 的联合密度. ($q(u, v) = 0, \forall (u, v) \notin G$).

$$\begin{aligned} P((U, V) \in D) &= P((X, Y) \in D^*) = P((X, Y) \in D^* \cap A) \\ &= \iint_{D^* \cap A} p(x, y) dx dy = \iint_{D \cap G} q(u, v) du dv = \iint_D q(u, v) du dv. \end{aligned}$$

例2.5. 设 X, Y 相互独立, 都服从 $N(0, 1)$. 极坐标表达:

$$X = R \cos \Theta, \quad Y = R \sin \Theta, \quad (R \geq 0, 0 \leq \Theta < 2\pi).$$

求 (R, Θ) 的联合密度与边缘密度.

- 解: 目标: 建立一、一映射. 解决: 回避 (x, y) 的坐标轴.
- 故, 取 $A = \{(u, v) : u \neq 0, v \neq 0\}$.
相应地, $G = \{(r, \theta) : r > 0, 0 < \theta < 2\pi \text{ 但 } \theta \neq \frac{1}{2}\pi, \pi, \frac{3}{2}\pi\}$.
- (R, Θ) 的联合密度: $(r, \theta) \in G$,

$$q(r, \theta) = p(x, y) \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} r = \frac{1}{2\pi} r e^{-\frac{1}{2}r^2}.$$

- 改变 $q(r, \theta)$ 在三条线上的值, 得到(新的)联合密度:

$$\varphi(r, \theta) = \frac{1}{2\pi} r e^{-\frac{1}{2}r^2}, \quad (r > 0, 0 < \theta < 2\pi).$$

- 注: R 服从瑞利分布, $\Theta \sim U[0, 2\pi]$, (且它们相互独立).

例2.4. 设 U, V 独立, 都服从 $U[0, 1]$. 求如下 (X, Y) 的联合密度.

$$X = R \cos \Theta, Y = R \sin \Theta; \quad \text{其中, } R = \sqrt{-2 \ln U}, \Theta = 2\pi V.$$

- 解: $r^2 = -2 \ln u, \theta = 2\pi v$.
- 故, 取 $A = \{(u, v) : 0 < u, v < 1, \text{但 } v \neq \frac{1}{4}, \frac{2}{4}, \frac{3}{4}\}$.
相应地, $G = \{(x, y) : x \neq 0, y \neq 0\}$.
- 雅可比行列式:

$$\frac{\partial(x, y)}{\partial(r, \theta)} = r, \quad \frac{\partial(r, \theta)}{\partial(u, v)} = \begin{vmatrix} r'(u) & 0 \\ 0 & \theta'(v) \end{vmatrix} = \frac{(r^2)'(u)}{2r} \cdot 2\pi = -\frac{2\pi}{ur}.$$

- 由连锁法则,

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial(u, v)}{\partial(r, \theta)} \cdot \frac{\partial(r, \theta)}{\partial(x, y)} = -\frac{ur}{2\pi} \cdot \frac{1}{r} = -\frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

- (U, V) 的联合密度: $p_{U,V}(u, v) = 1, (u, v) \in A$.
- 故, (X, Y) 的联合密度为

$$q(x, y) = p_{U,V}(u, v) \cdot \left| \frac{\partial(u, v)}{\partial(x, y)} \right| = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}, (x, y \neq 0).$$

- 改变 $q(x, y)$ 在坐标轴上的函数值, 得到(新的)联合密度:

$$p_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

- 注: X, Y 相互独立, 都服从 $N(0, 1)$.
- 注: 对比例2.5. 本质: R 服从瑞利分布当且仅当 $U \sim U(0, 1)$.

$$X = R \cos \Theta, Y = R \sin \Theta; \quad \text{其中, } R = \sqrt{-2 \ln U}, \Theta = 2\pi V.$$

例2.6. 设 U, V 相互独立, 都服从 $U(0, 1)$; $0 \leq r_1 < r_2$. 令

$$X = R \cos \Theta, \quad Y = R \sin \Theta,$$

$$\text{其中, } R = \sqrt{r_1^2 + (r_2^2 - r_1^2)U}, \quad \Theta = 2\pi V.$$

则 $(X, Y) \sim U(D)$, 其中 $D = \{(x, y) : r_1^2 \leq x^2 + y^2 \leq r_2^2\}$.

- 证: (R, Θ) 的取值区域: $A = \{(r, \theta) : 0 < r < 1, 0 < \theta < 1\}$, 相应地, (X, Y) 的取值区域: D . (去掉几条线.)
- 雅可比行列式:

$$\frac{\partial(x, y)}{\partial(r, \theta)} = r, \quad \frac{\partial(r, \theta)}{\partial(u, v)} = r'(u)\theta'(v)$$

$$\Rightarrow \frac{\partial(u, v)}{\partial(x, y)} = \frac{1}{r} u'(r) v'(\theta) = \frac{1}{r} \cdot \frac{2r}{r_2^2 - r_1^2} \cdot \frac{1}{2\pi}, \quad (u = \frac{r^2 - r_1^2}{r_2^2 - r_1^2}).$$

- $(X, Y) \sim U(D)$, 因为其联合密度如下:

$$p(x, y) = p_{U, V}(u, v) |\star| = \star, \quad (x, y) \in D.$$

两个随机变量的函数的均值公式

§4.3 随机向量的数字特征

- 设 (X, Y) 有密度 $p(x, y)$, 令 $Z = f(X, Y)$. 则

$$E(Z) = Ef(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)p(x, y)dxdy. \quad (3.1)$$

- 离散型类似.

例3.1. 设 X, Y 独立, 都服从 $N(0, 1)$, 求 $E\sqrt{X^2 + Y^2}$.

- 解法1: 直接用公式(3.1).

$$\begin{aligned} E\sqrt{X^2 + Y^2} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sqrt{x^2 + y^2} \cdot \frac{1}{2\pi} \cdot e^{-\frac{1}{2}(x^2+y^2)} dx dy \\ &= \int_0^{2\pi} d\theta \int_0^{\infty} r \cdot \frac{1}{2\pi} \cdot e^{-\frac{1}{2}r^2} r dr \quad (\text{作极坐标变换}) \\ &= \int_0^{\infty} r^2 e^{-\frac{1}{2}r^2} dr = \frac{\sqrt{2\pi}}{2} \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \frac{\sqrt{2\pi}}{2}. \end{aligned}$$

- 解法2: 先求 $Z = \sqrt{X^2 + Y^2}$ 的密度, 再按定义计算.

$$\begin{aligned} p_Z(z) &= ze^{-\frac{1}{2}z^2}, \quad (z > 0), \quad (\S 4.2, \text{例} 2.2) \\ \Rightarrow EZ &= \int_{-\infty}^{\infty} zp(z) dz = \int_0^{\infty} z^2 e^{-\frac{1}{2}z^2} dz = \frac{\sqrt{2\pi}}{2}. \end{aligned}$$

- 设 (X, Y) 的联合密度为 $p(x, y)$, 则

$$E(X) = \int_{-\infty}^{\infty} xp_X(x)dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xp(x, y)dxdy,$$

$$E(Y) = \int_{-\infty}^{\infty} yp_Y(y)dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yp(x, y)dxdy,$$

$$\begin{aligned} D(X) &= \int_{-\infty}^{\infty} (x - EX)^2 p_X(x)dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - EX)^2 p(x, y)dxdy, \end{aligned}$$

$$\begin{aligned} D(Y) &= \int_{-\infty}^{\infty} (y - EY)^2 p_Y(y)dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - EY)^2 p(x, y)dxdy. \end{aligned}$$

$$E(X + Y) = E(X) + E(Y). \quad (3.3)$$

- 证: 仅对 (X, Y) 为连续型的情形证明. 设联合密度为 $p(x, y)$.

$$\begin{aligned} E(X + Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y)p(x, y)dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xp(x, y)dx dy \\ &\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yp(x, y)dx dy \\ &= EX + EY. \end{aligned}$$

$$D(X + Y) = D(X) + D(Y) + 2E(X - EX)(Y - EY). \quad (3.4)$$

- 证: 记 $EX = a$, $EY = b$, 则 $E(X + Y) = a + b$.
- 记 $f(x, y) = ((x + y) - (a + b))^2$; $g(x, y) = (x - a)(y - b)$. 则

$$\begin{aligned} f(x, y) &= ((x - a) + (y - b))^2 \\ &= (x - a)^2 + (y - b)^2 + 2g(x, y). \end{aligned}$$

- $D(X + Y) = Ef(X, Y)$:

$$Ef(X, Y) = E(X - a)^2 + E(Y - b)^2 + 2Eg(X, Y).$$

若 X, Y 独立, 则

$$E(XY) = (EX) \cdot (EY). \quad (3.5)$$

- 证: 仅对连续型情形证明. 由独立性, 联合密度为

$$p(x, y) = p_X(x) \cdot p_Y(y)$$

- 于是

$$\begin{aligned} E(X \cdot Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyp(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyp_X(x)p_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} xp_X(x) dx \cdot \int_{-\infty}^{\infty} yp_Y(y) dy \\ &= (EX) \cdot (EY). \end{aligned}$$

若 X, Y 独立, 则

$$D(X + Y) = D(X) + D(Y). \quad (3.6)$$

$$E(X - EX)(Y - EY) = 0.$$

- 记 $EX = a, EY = b$, 则

$$(X - EX)(Y - EY) = (X - a)(Y - b) = XY - aY - bX + ab.$$

- 由独立性, $E(XY) = (EX) \cdot (EY) = ab$.

- 于是

$$\begin{aligned} E(X - EX)(Y - EY) &= E(XY - aY - bX + ab) \\ &= E(XY) - a \cdot EY - b \cdot EX + ab = ab - ab - ba + ab = 0. \end{aligned}$$

- 进一步, 再由(3.4) 即得(3.6).

- 定义3.1. 称实向量 (EX, EY) 为随机向量 (X, Y) 的期望/均值, 称实数 $E(X - EX)(Y - EY)$ 为 X, Y 的协方差, 记为 $\text{Cov}(X, Y)$ 或 σ_{XY} .
- $D(X) = \sigma_{XX}, D(Y) = \sigma_{YY}$.
- 设 (X, Y) 为连续型随机向量, 联合密度为 $p(x, y)$, 则

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - EX)(y - EY)p(x, y) dx dy.$$

- 若 X, Y 相互独立, 则 $\text{Cov}(X, Y) = 0$.
- 若 $\text{Cov}(X, Y) = 0$, 则称 X, Y (线性)不相关.
- 注: 独立 \Rightarrow 不相关, 不相关 \nRightarrow 独立.

例3.2 设 $(X, Y) \sim U(D)$, $D = \{(x, y) : x^2 + y^2 \leq 1\}$.

求 $\sigma_{XX}, \sigma_{YY}, \sigma_{XY}$.

● 解: $EX = EY = 0$:

$$EX = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xp(x, y) dx dy = \iint_{x^2+y^2 \leq 1} x \cdot \frac{1}{\pi} dx dy$$

$$= \frac{1}{\pi} \int_{-1}^1 \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy \right) x dx = 0.$$

● $\sigma_{XX} = \sigma_{YY} = \frac{1}{4}$:

$$\sigma_{XX} = \iint_{x^2+y^2 \leq 1} (x - 0)^2 \cdot \frac{1}{\pi} dx dy$$

$$= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 r^2 \cos^2 \theta \cdot r dr d\theta \quad (\text{作极坐标变换})$$

$$= \frac{1}{\pi} \left(\int_0^{2\pi} \cos^2 \theta d\theta \right) \cdot \left(\int_0^1 r^3 dr \right) = \frac{1}{\pi} \cdot \pi \cdot \frac{1}{4} = \frac{1}{4}.$$

- $\sigma_{XY} = 0$:

$$\begin{aligned}\sigma_{XY} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - EX)(y - EY)p(x, y)dx dy \\ &= \frac{1}{\pi} \iint_{x^2+y^2 \leq 1} xy dx dy \\ &= \frac{1}{\pi} \int_{-1}^1 \left(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x dx \right) y dy = 0.\end{aligned}$$

- 注: X 和 Y 不相关, 但是 X 和 Y 不独立.
- “ X 和 Y 不独立”的直观: 当 $X = x$ 时, Y 的取值范围为 $[-\sqrt{1-x^2}, \sqrt{1-x^2}]$, 它依赖于 x .
- “ X 和 Y 不独立”的直观: 取 $A = (\frac{\sqrt{2}}{2}, 1]$, $B = (\frac{\sqrt{2}}{2}, 1]$, 则 $P(X \in A) > 0$, $P(Y \in B) > 0$, 但

$$P(X \in A, Y \in B) = 0 \neq P(X \in A) \cdot P(Y \in B).$$

例3.3. 设 (X, Y) 服从二维正态分布, 密度函数如下. 求 σ_{XY} .

$$p(x, y) = C e^{-\frac{1}{2(1-\rho^2)} \cdot I}, \quad \text{其中, } C = 1/(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}),$$
$$I = u^2 - 2\rho uv + v^2. \quad u = (x - \mu_1)/\sigma_1, \quad v = (y - \mu_2)/\sigma_2.$$

- 解: 已有结论: $EX = \mu_1, EY = \mu_2, D(X) = \sigma_1^2, D(Y) = \sigma_2^2$.

$$\begin{aligned}\sigma_{XY} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_1)(y - \mu_2) p(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma_1 u \cdot \sigma_2 v \cdot C e^{-\frac{1}{2(1-\rho^2)} \cdot I} \sigma_1 \sigma_2 du dv \\ &= \frac{\sigma_1 \sigma_2}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} u \cdot e^{-\frac{(u-\rho v)^2 + (1-\rho^2)v^2}{2(1-\rho^2)}} du \right) v dv \\ &= \frac{\sigma_1 \sigma_2}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \left(\sqrt{2\pi}\sqrt{1-\rho^2} \cdot \rho v \cdot e^{-\frac{v^2}{2}} \right) v dv \\ &= \frac{\sigma_1 \sigma_2 \rho}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v^2 e^{-\frac{v^2}{2}} dv = \sigma_1 \sigma_2 \rho.\end{aligned}$$

- 注: 对于二维正态分布的两个分量, 不相关 \Rightarrow 独立.

相关系数

- 定义3.2. 称

$$\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}\sigma_{YY}}}$$

为 X, Y 的相关系数, 记作 ρ_{XY} 或 ρ .

- 注: 要求 $0 < \sigma_{XX}, \sigma_{YY} < \infty$.
- 注: 对二元正态分布, 参数 ρ 是两个分量的相关系数.
- 注: 将 X, Y 分别标准化, 得到 $X^* = (X - EX)/\sqrt{\sigma_{XX}}$,
 $Y^* = (Y - EY)/\sqrt{\sigma_{YY}}$. 则 $\rho_{XY} = \sigma_{X^*Y^*}$:

$$\rho_{XY} = \frac{E(X - EX)(Y - EY)}{\sqrt{\sigma_{XX}\sigma_{YY}}} = EX^*Y^* = \sigma_{X^*Y^*}.$$

相关系数 $\rho = \rho_{XY}$ 满足:

$$|\rho| \leq 1. \quad (3.10)$$

- 证: 考虑一元二次函数 $f: \lambda \rightarrow f(\lambda)$, 定义如下:

$$\begin{aligned} f(\lambda) &:= D(Y - \lambda X) \\ &= D(Y) + D(-\lambda X) + 2E(Y - EY)(-\lambda X - E(-\lambda X)) \\ &= \sigma_{YY} + \lambda^2 \sigma_{XX} - 2\lambda \sigma_{XY} = \sigma_{XX} \lambda^2 - 2\sigma_{XY} \lambda + \sigma_{YY}. \end{aligned}$$

- f 的最小值点: $\lambda_0 = \frac{\sigma_{XY}}{\sigma_{XX}}$, 最小值: $f(\lambda_0) = \sigma_{YY}(1 - \rho^2)$:

$$f(\lambda) = \sigma_{XX} \left(\lambda - \frac{\sigma_{XY}}{\sigma_{XX}} \right)^2 + \sigma_{YY} - \frac{\sigma_{XY}^2}{\sigma_{XX}}.$$

- 方差总是非负的, 因此 $1 - \rho^2 \geq 0$, 即 $|\rho| \leq 1$.
- 注: $|\rho| = 1$ 当且仅当 $D(Y - \lambda_0 X) = 0$,
即存在常数 a, b 使得 $P(Y = a + bX) = 1$.

线性预测

- 预测问题是统计学的重要研究课题.
- 线性函数 $f(x) = a + bx$ 是“最简单”的函数.
- 下设 $\sigma_{XX} > 0, \sigma_{YY} > 0$.
问题: 如何选取 a, b , 使得 $a + bX$ 与 Y 最接近?
- 用预测的均方误差 $Q(a, b)$ 刻画接近程度:

$$Q = Q(a, b) = E(Y - (a + bX))^2.$$

- 问题: 求 $Q(a, b)$ 的最小值点 (a^*, b^*) .
- 解: $Q(a, b)$ 是关于 a, b 的二元二次多项式. 可用配方的办法或偏导数等于零的办法来求最小值点.

计算 $Q(a, b) = E(Y - (a + bX))^2$.

- $Y - (a + bX) = (Y - EY) - b(X - EX) + EY - bEX - a$.
- $Z = (Y - EY) - b(X - EX)$ 满足 $EZ = 0$;
 $c = EY - bEX - a$ 为常数/实数.
- 展开:

$$Q(a, b) = E(Z + c)^2 = EZ^2 + c^2 + 2cEZ = EZ^2 + c^2.$$

- 进一步, 将 Z^2 再展开:

$$\begin{aligned} EZ^2 &= \sigma_{YY} + b^2\sigma_{XX} - 2b\sigma_{XY} \\ &= \sigma_{XX}b^2 - 2\sigma_{XY}b + \sigma_{YY}. \end{aligned}$$

配方方法:

- $Q(a, b) = E(Y - (a + bX))^2 = EZ^2 + c^2$, 其中

$$EZ^2 = \sigma_{XX}b^2 - 2\sigma_{XY}b + \sigma_{YY}, \quad c = EY - bEX - a.$$

- $EZ^2 = \sigma_{XX} \left(b - \frac{\sigma_{XY}}{\sigma_{XX}} \right)^2 + \sigma_{YY} - \frac{\sigma_{XY}^2}{\sigma_{XX}}$.
- 目标: 使 $Q(a, b)$ 达到最小.
- 取 $b^* = \frac{\sigma_{XY}}{\sigma_{XX}} = \frac{\rho\sqrt{\sigma_{XX}\sigma_{YY}}}{\sigma_{XX}} = \rho\sqrt{\frac{\sigma_{YY}}{\sigma_{XX}}}$ 使得 EZ^2 达到最小.
取 $a^* = EY - b^*EX$ 使得 $c = 0$.
- 二元函数 $Q(a, b)$ 的最小值:

$$Q(a^*, b^*) = \sigma_{YY} - \frac{\sigma_{XY}^2}{\sigma_{XX}} = \sigma_{YY}(1 - \rho^2).$$

- 注: $|\rho|$ 越接近于1, \star 越小.
- 注: $\rho \neq 0$, 则 $\star < \sigma_{YY} = E(Y - EY)^2 = Q(EY, 0)$.

偏导法:

- $Q(a, b) = E(Y - (a + bX))^2 = EZ^2 + c^2$, 其中

$$EZ^2 = \sigma_{XX}b^2 - 2\sigma_{XY}b + \sigma_{YY}, \quad c = EY - bEX - a.$$

- 目标: 使 $Q(a, b)$ 达到最小.
- 求解方程组:

$$\begin{cases} \frac{\partial Q(a, b)}{\partial a} = 0, \\ \frac{\partial Q(a, b)}{\partial b} = 0. \end{cases} \quad \text{即, } \begin{cases} -2c = 0, \\ 2\sigma_{XX}b - 2\sigma_{XY} + 2c(-EX) = 0. \end{cases}$$

解得,
$$\begin{cases} a^* = EY - bEX, \\ b^* = \frac{\sigma_{XY}}{\sigma_{XX}}. \end{cases}$$

- 检查 (a^*, b^*) 确实是最小解.

§4.4 关于 n 维随机向量

- n 维随机向量 $\xi = (X_1, X_2, \dots, X_n)$.
- 定义4.1. 若存在非负函数 $p(x_1, \dots, x_n)$ 使得:

$$P(\xi \in D) = \int \cdots \int_D p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \quad (4.1)$$

对任意 n 维长方体

$$D = \{(x_1, x_2, \dots, x_n) : a_1 < x_1 < b_1, \dots, a_n < x_n < b_n\}$$

均成立, 则称 ξ 是连续型的, 称 $p(x_1, x_2, \dots, x_n)$ 为 ξ 的分布密度, 或联合分布密度, 简称联合密度.

- 注: (4.1) 对于 n 维空间中的区域, 甚至相当一般的集合 D 仍成立.
- 称 $\xi = (X_1, \dots, X_n)$ 的一部分分量所构成的向量及其分布密度 **边缘** 及其 **边缘密度**. 如, (X_2, X_5) 为 ξ 的二维边缘, $p_{25}(x, y) := p_{X_2, X_5}(x, y)$ 为二维边缘密度.
- 特别地, 每个分量 X_i 的分布密度 $p_i(x)$ 是 ξ 一维边缘密度, 称为单个密度.
- 联合密度决定边缘密度: 如,

$$p_1(\mathbf{x}_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) dx_2 \dots dx_n, \quad (4.3)$$

$$p_{12}(\mathbf{x}_1, \mathbf{x}_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n) dx_3 \dots dx_n.$$

- 定义4.2. 若 $\xi = (X_1, \dots, X_n)$ 的联合密度 $p(x_1, \dots, x_n)$ 具有如下表达式, 则称 ξ 服从 n 维正态分布, 称 ξ 是 n 维正态随机向量, 记为 $\xi \sim N(\mu, \Sigma)$.

$$p(x_1, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\} \quad (4.4)$$

其中 $x = (x_1, \dots, x_n)^T$ (自变量), $\mu = (\mu_1, \dots, \mu_n)$ 是 n 维向量, $\Sigma = (\sigma_{ij})_{n \times n}$ 是 n 阶正定矩阵.

独立性

- 定义4.3. 设 X_1, \dots, X_n 是 n 个随机变量. 如果对任意 $a_i < b_i$ ($i = 1, 2, \dots, n$), 事件 $\{a_1 < X_1 < b_1\}, \dots, \{a_n < X_n < b_n\}$ 相互独立, 则称 X_1, \dots, X_n 是相互独立的.
- 定理4.1 设 X_1, \dots, X_n 的分布密度分别是 $p_1(x_1), \dots, p_n(x_n)$, 则 X_1, \dots, X_n 相互独立的充分必要条件是

$$p(x_1, \dots, x_n) := p_1(x_1) \dots p_n(x_n)$$

是 (X_1, \dots, X_n) 的联合分布密度.

- 例, 若 X_1, X_2, X_3 相互独立, 则 $\{(X_1, X_2) \in D\}$ 与 $\{X_3 \in B\}$ 独立.

- 定义: 若 $\xi = (X_1, \dots, X_n)$ 与 $\eta = (Y_1, \dots, Y_m)$ 满足:
对任意 $a_i < b_i$ ($i = 1, 2, \dots, n$); $c_j < d_j$ ($j = 1, 2, \dots, m$),
 $\{a_1 < X_1 < b_1, \dots, a_n < X_n < b_n\}$
与 $\{c_1 < Y_1 < d_1, \dots, c_m < Y_m < d_m\}$ 相互独立,
则称 ξ 与 η 相互独立.

- 设 ξ, η 都是连续型随机向量, 联合密度分别为 $p_\xi(x_1, \dots, x_n)$,
 $p_\eta(y_1, \dots, y_m)$. 则 ξ 与 η 相互独立的充分必要条件是:

$$q(x_1, \dots, x_n, y_1, \dots, y_m) := p_\xi(x_1, \dots, x_n)p_\eta(y_1, \dots, y_m)$$

为 $(X_1, \dots, X_n, Y_1, \dots, Y_m)$ 的联合密度.

- 注: 类似地, 可以定义 n 个随机向量相互独立.

n 个随机变量的函数的分布

- 设 (X_1, \dots, X_n) 为连续型, 联合密度为 $p(x_1, \dots, x_n)$.
- 设 $Y = f(X_1, X_2, \dots, X_n)$. 则 Y 的分布函数为

$$\begin{aligned} F_Y(y) &= P(f(X_1, \dots, X_n) \leq y) \\ &= \int \cdots \int_{f(x_1, \dots, x_n) \leq y} p(x_1, \dots, x_n) dx_1 \cdots dx_n. \end{aligned} \quad (4.5)$$

- 注: 设 $\xi = (X_1, \dots, X_n)$ 与 $\eta = (Y_1, Y_2, \dots, Y_m)$ 相互独立, 则 $f(X_1, \dots, X_n)$ 与 $g(Y_1, \dots, Y_m)$ 相互独立.

- 连续型随机向量的函数的期望, 计算公式:

$$Ef(X_1, \dots, X_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) p(x_1, \dots, x_n) dx_1 \dots dx_n.$$

注: 要求右端的积分绝对收敛.

- 对于离散型随机向量, 积分变成求和.
- 称实向量 (EX_1, \dots, EX_n) 为随机向量 (X_1, \dots, X_n) 的期望/均值.

- 和的期望:

$$E(X_1 + \cdots + X_n) = EX_1 + \cdots + EX_n.$$

- 若 X_1, X_2, \dots, X_n 相互独立, 则

$$E(X_1 \cdots X_n) = (EX_1)(EX_2) \cdots (EX_n), \quad (4.8)$$

$$D(X_1 + \cdots + X_n) = D(X_1) + \cdots + D(X_n).$$

- 证: $n = 2, \checkmark$. 数学归纳法.

- $\xi = (X_1, \dots, X_n)$ 是 n 维随机向量.
- 对 $i, j = 1, 2, \dots, n$, 记

$$\sigma_{ij} = \sigma_{X_i, X_j} = E(X_i - EX_i)(X_j - EX_j). \quad (4.9)$$

- 称矩阵

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

为 (X_1, X_2, \dots, X_n) 的协方差矩阵, 简称协方差阵或协差阵.

- 注: Σ 是实对称矩阵.

Σ 非负定的/半正定的.

- 半正定矩阵的定义:

对任意实数 a_1, \dots, a_n , 令 $a = (a_1, \dots, a_n)^T$, 则 $a^T \Sigma a \geq 0$.

- 证:

$$\begin{aligned} a^T \Sigma a &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \sigma_{ij} \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j E(X_i - EX_i)(X_j - EX_j) \\ &= E \left(\sum_{i=1}^n a_i (X_i - EX_i) \cdot \sum_{j=1}^n a_j (X_j - EX_j) \right) \\ &= E \left(\sum_{i=1}^n a_i (X_i - EX_i) \right)^2 = D \left(\sum_{i=1}^n a_i X_i \right) \geq 0. \end{aligned}$$

- 进一步, 记

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}.$$

注: $\rho_{ii} = 1$.

- 称矩阵

$$R = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{pmatrix}$$

为 (X_1, \dots, X_n) 的相关系数阵, 简称相关阵.

- 注: R 也是实对称、也是非负定的.

- 记

$$C = \begin{pmatrix} \frac{1}{\sqrt{\sigma_{11}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{22}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{1}{\sqrt{\sigma_{nn}}} \end{pmatrix}$$

则

$$R = C\Sigma C.$$

- 注: 将 X_i 标准化, $Y_i = (X_i - EX_i)/\sqrt{\sigma_{ii}}$, 则 R 是 (Y_1, \dots, Y_n) 的协方差阵.

- 定义4.4 设 $\xi = (X_1, \dots, X_n)$ 是 n 维随机向量, 称 n 元函数

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

为 ξ 的分布函数或联合分布函数.

- 如果 ξ 有联合密度 $p(x_1, \dots, x_n)$, 则

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(u_1, \dots, u_n) du_1 \cdots du_n.$$

- 注: 在联合密度的连续点上,

$$p(x_1, \dots, x_n) = \frac{\partial F(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n}.$$

§4.5 条件分布与条件期望

- 设 (X, Y) 是离散型随机向量. 联合分布(列)为

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

- 固定 y_j . 称

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

为在 $Y = y_j$ 的条件下, X 的条件分布(列).

- 在 $X = x_i$ 的条件下, Y 的条件分布(列)类似定义.
- 乘法公式: $p_{ij} = P(Y = y_j) \cdot P(X = x_i | Y = y_j)$.
- 注: 若 X 的条件分布(列)不依赖于 y_j , 则 X 与 Y 相互独立.

例5.1. 射手单发击中目标的概率: p , ($0 < p < 1$), 击中两次时停.

$X =$ 第一次击中目标时的射击次数; $Y =$ 总射击次数.

求联合分布、边缘分布、条件分布.

- 解: 记 $q = 1 - p$.

联合分布: $n > m \geq 1$,

$$P(X = m, Y = n) = q^{m-1}p \cdot q^{n-m-1}p = q^{n-2}p^2.$$

- 边缘分布:

$$P(X = m) = \sum_{n=m+1}^{\infty} q^{n-2}p^2 = q^{m-1}p, \quad m \geq 1;$$

$$P(Y = n) = \sum_{m=1}^{n-1} q^{n-2}p^2 = (n-1)q^{n-2}p^2, \quad n \geq 2.$$

例5.1(续).

- 在 $Y = n$ 的条件下, X 的分布: $m = 1, \dots, n - 1$,

$$P(X = m|Y = n) = \frac{P(X = m, Y = n)}{P(Y = n)} = \frac{q^{n-2}p^2}{(n-1)q^{n-2}p^2} = \frac{1}{n-1}.$$

- 在 $X = m$ 的条件下, Y 的条件分布: $n = m + 1, m + 2, \dots$,

$$P(Y = n|X = m) = \frac{P(X = m, Y = n)}{P(X = m)} = \frac{q^{n-2}p^2}{q^{m-1}p} = q^{n-m-1}p.$$

- 注: 等价地, 对于 $k = 1, 2, \dots$

$$P(Y - X = k|X = m) = P(Y = m + k|X = m) = q^{k-1}p.$$

此即 X 的分布列. 于是, $Y - X$ 与 X 独立且同分布.

连续型条件分布

- 设 (X, Y) 为连续型随机向量, 联合密度为 $p(x, y)$.
- Y 的边缘密度: $p_Y(y) = \int_{-\infty}^{\infty} p(x, y)dx$.
- 固定 y . 设 $p_Y(y) > 0$. 称关于 x 的函数

$$p_{X|Y}(x|y) := \frac{p(x, y)}{p_Y(y)} \quad (5.4)$$

为在 $Y = y$ 的条件下, X 的条件(分布)密度.

- 在 $X = x$ 的条件下, Y 的条件分布密度类似定义.
- 注: 例, 若 $p_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, 则称在 $Y = y$ 的条件下, X 服从 $N(0, 1)$.
- 乘法公式: $p(x, y) = p_Y(y) \cdot p_{X|Y}(x|y)$.
- 注: 若 X 的条件分布密度不依赖于 y , 则 X 与 Y 相互独立.

例5.2. 设 (X, Y) 服从二维正态分布, 密度函数如下. 求 $p_{X|Y}(x|y)$.

$$p(x, y) = C e^{-\frac{1}{2(1-\rho^2)} \cdot I}, \quad \text{其中, } C = 1/(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}),$$
$$I = u^2 - 2\rho uv + v^2. \quad u = (x - \mu_1)/\sigma_1, \quad v = (y - \mu_2)/\sigma_2.$$

- 方法一、先求 $p_Y(y)$. 已有结论: $Y \sim N(\mu_2, \sigma_2^2)$, 即

$$p_Y(y) = C_1 e^{-\frac{v^2}{2}}, \quad \text{其中 } C_1 = 1/(\sqrt{2\pi}\sigma_2).$$

- 在 $Y = y$ 的条件下, X 的条件分布密度:

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \frac{C}{C_1} e^{-\frac{1}{2(1-\rho^2)} \cdot I} e^{\frac{v^2}{2}}$$
$$= \frac{C}{C_1} e^{-\frac{(u-\rho v)^2}{2(1-\rho^2)}} e^{-\frac{(1-\rho^2)v^2}{2(1-\rho^2)}} e^{\frac{v^2}{2}} = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_1} e^{-\frac{(x-\mu_1-\sigma_1\rho v)^2}{2(1-\rho^2)\sigma_1^2}}.$$

- 在 $Y = y$ 的条件下, X 服从 $N(\mu, \sigma^2)$, 其中

$$\mu = \mu_1 + \rho\sigma_1 \frac{y-\mu_2}{\sigma_2}, \quad \sigma^2 = (1-\rho^2)\sigma_1^2.$$

$$p(x, y) = Ce^{-\frac{1}{2(1-\rho^2)} \cdot I}, \quad \text{其中, } C = 1/(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}),$$

$$I = u^2 - 2\rho uv + v^2. \quad u = (x - \mu_1)/\sigma_1, \quad v = (y - \mu_2)/\sigma_2.$$

- 方法二、直接从联合密度出发:

$$p(x, y) = Ce^{-\frac{1}{2(1-\rho^2)} \cdot I} = Ce^{-\frac{(u-\rho v)^2}{2(1-\rho^2)}} e^{-\frac{(1-\rho^2)v^2}{2(1-\rho^2)}}$$

$$= C_y e^{-\frac{(x-\mu_1-\sigma_1\rho v)^2}{2(1-\rho^2)\sigma_1^2}}.$$

- 记 $\mu = \mu_1 + \rho\sigma_1 \frac{y-\mu_2}{\sigma_2}$, $\sigma^2 = (1-\rho^2)\sigma_1^2$. 则

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \hat{C}_y e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (\hat{C}_y = \frac{C_y}{p_Y(y)}).$$

- 上式关于 x 是密度, 故必有 $\hat{C}_y = \frac{1}{\sqrt{2\pi\sigma^2}}$.
- 注: 类似地, 在 $X = x$ 的条件下, Y 服从 $N(\mu, \sigma^2)$, 其中

$$\mu = \mu_2 + \rho\sigma_2 \frac{x-\mu_1}{\sigma_1}, \quad \sigma^2 = (1-\rho^2)\sigma_2^2.$$

- 定义5.2. 设 (X, Y) 是连续型随机向量, 称

$$\int_{-\infty}^{\infty} xp_{X|Y}(x|y)dx$$

为在 $Y = y$ 的条件下, X 的条件期望, 即为 $E(X|Y = y)$.

- 注: 要求上面的积分绝对收敛.
- 对离散型, 可类似定义条件期望.
- 注: 条件期望即为条件分布对应的期望.

- 根据条件密度的定义,

$$g(y) := E(X|Y = y) = \frac{1}{p_Y(y)} \int_{-\infty}^{\infty} xp(x, y)dx. \quad (5.5)$$

- $g(y)$ 是 y 的函数. $g(Y)$ 是随机变量, 记为 $E(X|Y)$, 它是随机变量 Y 的函数.
- 重期望公式:

$$\begin{aligned} EE(X|Y) &= Eg(Y) = \int_{\{y:p_Y(y)>0\}} g(y)p_Y(y)dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xp(x, y)dx dy = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} p(x, y)dy dx \\ &= \int_{-\infty}^{\infty} xp_X(x)dx = EX. \end{aligned} \quad (5.6)$$

- 注: 本质是全概公式.

例5.3. 设 U_1, U_2, \dots 相互独立, 都服从 $U(0, 1)$. 求 EN , 其中
 $N := \min \{n : \sum_{i=1}^n U_i > 1\}$.

- 解: 对任意 $x \in [0, 1]$, 令

$$N(x) := \min \{n : \sum_{i=1}^n U_i > x\}, \quad m(x) = EN(x).$$

- 则 $g(y) := E(N(x)|U_1 = y) = \begin{cases} 1, & y > x, \\ 1 + m(x - y), & y \leq x. \end{cases}$

- 故

$$\begin{aligned} m(x) &= Eg(U_1) = \int_x^1 dy + \int_0^x (1 + m(x - y)) dy \\ &= 1 + \int_0^x m(u) du. \end{aligned}$$

- 等价地, $m'(x) = m(x)$. 解得 $m(x) = ke^x$.
- 由 $m(0) = 1$ 知 $k = 1$. 故 $m(x) = e^x$. 所求为 $EN = m(1) = e$.

例5.4. 设某工厂每月电力需求: $X \sim U(10, 20)$, 电力供应:
 $Y \sim U(10, 30)$, 单位: 万度. 若电力供应足够, 每1万度电可创
 造30万元利润; 若不足, 则不足部分电力每1 万度创造10万元利
 润. 求此工厂每月平均利润.

● 解: 可认为 X, Y 独立. 工厂每月的利润 R (万元)表达式如下:
 $R = 30X$, (若 $X \leq Y$); $R = 30Y + 10(X - Y)$, (若 $X > Y$).

● 记 $g(y) = E(R|Y = y)$. 当 $20 \leq y \leq 30$ 时, 必有 $X \leq y$,

$$g(y) = E(30X) = 30EX = 30 \cdot \frac{10+20}{2} = 450.$$

● 当 $10 \leq y < 20$ 时,

$$g(y) = \int_{10}^y 30x \cdot \frac{1}{10} dx + \int_y^{20} (30y + 10(x - y)) \cdot \frac{1}{10} dx = 50 + 40y - y^2.$$

● 答案: 约433 万元.

$$E(R) = g(Y) = \int_{10}^{20} (50 + 40y - y^2) \cdot \frac{1}{20} dy + \int_{20}^{30} 450 \cdot \frac{1}{20} dy \approx 433.$$

随机向量的条件分布和条件期望

- 设 $X = (X_1, \dots, X_m)$ 和 $Y = (Y_1, \dots, Y_n)$ 是两个随机向量。
设 (X, Y) 有联合密度 $p(x, y)$,
其中 $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_n)$.
- 称 $p_{X|Y}(x|y) = p(x, y)/p_Y(y)$ 为在 $Y = y$ 的条件下, X 的条件(联合)密度.
- 在 $Y = y$ 的条件下, X_i 的条件密度: $p_{X_i, Y}(x_i, y)/p_Y(y)$ 即为 $p_{X|Y}(x|y)$ 的第 i 个边缘密度.
- 在 $Y = y$ 条件下, X_i 的条件期望:

$$E(X_i|Y = y) := \int_{-\infty}^{\infty} u p_i(u|y) du.$$

- 在 $Y = y$ 条件下, X 的条件期望:

$$E(X|Y = y) := (E(X_1|Y = y), \dots, E(X_m|Y = y)).$$

最佳预测

- 求函数 $\psi(\cdot)$, 使得均方误差 $Q(\psi) := E(Y - \psi(X))^2$ 最小.
- 定理5.1. 设 (X, Y) 有联合密度 $p(x, y)$, $EY^2 < \infty$. 则

$$\phi(x) = E(Y|X = x), (\text{若 } p_X(x) > 0); \phi(x) = 0, (\text{若 } p_X(x) = 0).$$

为 $Q(\cdot)$ 的最小值点. (注: 离散型类似.)

- 证: $Y - \psi(X) = Y - \phi(X) + \phi(X) - \psi(X)$.
- $E\star\star = 0$:

$$E\star\star = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \phi(x))(\phi(x) - \psi(x))p_X(x)p_{Y|X}(y|x)dydx,$$

其中, $\int_{-\infty}^{\infty} (y - \phi(x))p_{Y|X}(y|x)dy = \int_{-\infty}^{\infty} yp_{Y|X}(y|x)dy - \phi(x) = 0$.

- $Q(\psi) = Q(\phi) + E(\phi(X) - \psi(X))^2$.

例5.5. 设 (X, Y) 服从二维正态分布, 参数为 $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$.

- 在 $X = x$ 的条件下, Y 服从 $N\left(\mu_2 + \rho\sigma_2\frac{x-\mu_1}{\sigma_1}, \sigma_2^2(1-\rho^2)\right)$.
- 条件期望

$$\phi(x) = E(Y|X = x) = \mu_2 + \rho\sigma_2\frac{x - \mu_1}{\sigma_1}.$$

- 在 X 的函数预测 Y , 使得均方误差最小的是

$$\phi(X) = \mu_2 + \rho\sigma_2\frac{X - \mu_1}{\sigma_1}.$$

多元最佳预测:

- 设 X_1, \dots, X_m, Y 是 $m + 1$ 个随机变量.
记 $X = (X_1, \dots, X_m)$, $x = (x_1, \dots, x_m)$.
- 求函数 $\phi(x)$, 使得用 $\phi(X)$ 预测 Y 的均方误差最小.
- 答案为 $\phi(x) = E(Y|X = x)$. 将 $\phi(X)$ 记为 $E(Y|X)$.

§4.6 大数定律和中心极限定理

- 随机变量序列: X_1, X_2, \dots
- 定义6.1. 若 $\forall n \geq 1, X_1, \dots, X_n$ 相互独立, 则称 X_1, X_2, \dots 是**相互独立的**. 又若所有的 X_i 具有相同的分布, 则称 X_1, X_2, \dots 是**独立同分布**的随机变量序列, 记为**i.i.d.**.
- 令

$$S_n = X_1 + \dots + X_n.$$

大数定律: 定理6.1. 设 X_1, X_2, \dots 独立同分布, 且 $EX_1, D(X_1)$ 存在. 则 $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{S_n}{n} - EX_1 \right| \geq \varepsilon \right) = 0. \quad (6.1)$$

- 注: $\bar{X} = \frac{S_n}{n}$ 是 n 个观测的算术平均值.
- 注: 令 $X_n = 1$, 若第 n 次试验中 A 发生; $X_1 = 0$, 否则. 那么, \bar{X} 为频率, $EX_1 = P(A)$. 频率 \approx 概率. (P155, 例6.1)
- 证: $E\bar{X} = EX_1$. 由切比雪夫不等式, $\star \leq \frac{1}{\varepsilon^2} D(\bar{X})$.
- 独立, 故 $D(S_n) = D(X_1) + \dots + D(X_n) = nD(X)$. 从而, $D(\bar{X}) = \frac{1}{n^2} D(S_n) = \frac{1}{n} D(X_1)$.
- 因此 $\star \leq \frac{D(X_1)}{n\varepsilon^2} \rightarrow 0, n \rightarrow \infty$.
- 注: 若(6.1) 成立, 则称 X_1, X_2, \dots 服从(弱)大数定律.

- 强大数定律: 设 X_1, X_2, \dots 独立同分布, 且 EX_1 存在, 则

$$P\left(\lim_{n \rightarrow \infty} \frac{S_n}{n} = EX_1\right) = 1. \quad (6.2)$$

- 注: 若(6.2) 成立, 则称 X_1, X_2, \dots 服从强大数定律.
- 设 ξ, ξ_1, ξ_2, \dots 均为随机变量. 若

$$\lim_{n \rightarrow \infty} P(|\xi_n - \xi| > \varepsilon) = 0, \quad \forall \varepsilon > 0,$$

则称 ξ_n 依概率收敛到 ξ , 记为 $\xi_n \xrightarrow{\text{Pr}} \xi$. 若

$$P\left(\lim_{n \rightarrow \infty} \xi_n = \xi\right) = 1,$$

则称 ξ_n 以概率1收敛到 ξ , 或几乎必然(a.s.)收敛到 ξ , 记为 $\xi_n \xrightarrow{\text{a.s.}} \xi$ 或 $\lim_{n \rightarrow \infty} \xi_n = \xi, \text{ a.s.}$.

- 注: 可以证明 $\xi_n \xrightarrow{\text{a.s.}} \xi$ 蕴含着 $\xi_n \xrightarrow{\text{Pr}} \xi$, 故(6.2) 蕴含着(6.1).

中心极限定理

- 中心极限定理: 定理6.2. 设 X_1, X_2, \dots 独立同分布;
 $\mu = EX_1, \sigma^2 = DX_1$ 存在; 且 $\sigma^2 > 0$. 则 $\forall a < b$,

$$\lim_{n \rightarrow \infty} P \left(a < \frac{S_n - nEX_1}{\sqrt{nD(X_1)}} < b \right) = \int_a^b \phi(u) du, \quad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

- 考虑标准化: $\xi^* = \frac{\xi - E\xi}{\sqrt{D(\xi)}}$. 则 $\star = S_n^*$.
- 固定 n . $\bar{X} = \frac{S_n}{n}$, 则

$$(\bar{X})^* = \frac{\frac{S_n}{n} - EX_1}{\sqrt{\frac{1}{n}D(X_1)}} = S_n^* = \star.$$

- 当 n 很大时, $S_n^* = (\bar{X})^*$ 近似服从 $N(0, 1)$. S_n 近似服从 $N(nEX_1, nD(X_1))$, \bar{X} 近似服从 $N(EX_1, \frac{D(X_1)}{n})$.
- 注: 中心极限定理是现代统计推断中一个重要基础理论.

例6.2. 某座桥的最大负荷重量 $Y \sim N(300, 40)$ (单位: 吨); 每辆车的平均重量为5, 方差为2. 为了保证桥不被压塌的概率不小于0.99997, 最多允许多少辆车在桥上?

- 解: 若有 M 辆车在桥上, 第 i 辆车重量为 X_i . M 辆车的总重量为 $S_n = X_1 + \cdots + X_M$.
- 可以认为 Y, X_1, X_2, \dots, X_M 相互独立;
 $\mu = EX_1 = 5, \sigma^2 = D(X_1) = 2$.
- 求最大的 M , 使得 $P(S_M < Y) \geq 0.99997$.
- M 较大, 故 S_M 近似服从 $N(M\mu, M\sigma^2)$.
- S_M 与 Y 独立, 都(近似)服从正态分布, 故 $Z = S_M - Y$ (近似)服从正态分布. $EZ = M\mu - 300, D(Z) = M\sigma^2 + 40$.
- 为使 $P(Z < 0) \geq 0.99997$, 即

$$P(Z < 0) \approx \Phi\left(\frac{0 - (M\mu - 300)}{\sqrt{M\sigma^2 + 40}}\right) \geq 0.99997.$$

例6.2(续).

- 等价地,

$$\frac{0 - (M\mu - 300)}{\sqrt{M\sigma^2 + 40}} \geq \Phi^{-1}(0.99997) = 4. \quad (\text{查表})$$

- 求解

$$\Leftrightarrow \frac{-(5M - 300)}{\sqrt{2M + 40}} \geq 4 \Leftrightarrow 300 - 5M \geq 4\sqrt{2M + 40}$$

$$\Leftrightarrow \begin{cases} 300 - 5M \geq 0 \\ 25M^2 - 3032M + 89360 \geq 0 \end{cases} \Leftrightarrow \begin{cases} M \leq 60 \\ M \leq 50.5 \text{ 或 } M \geq 70.78. \end{cases}$$

- 解得 $M \leq 50.5$. 故, 最多允许50 辆车同时在桥上.