Chap 3 联合分布

1 随机向量

- **定义(随机向量)** 我们称

$$(X_1, \cdots, X_n) : \Omega \to \mathbb{R}^n$$

为随机向量, 当 $X_i(1 \leq i \leq n)$ 均为随机变量.

- **定义(联合 CDF)**

$$F(x_1, \cdots, x_n) := F(X_1 \leq x_1, \cdots, X_n \leq x_n), \forall (x_1, \cdots, x_n) \in \mathbb{R}^n.$$

- **注** 若 $X_i : \Omega_i \to \mathbb{R}$, 需扩充 $\Omega = \Omega_1 \times \cdots \Omega_n$.

2 离散分布

- **定义(离散型随机向量)**

$$(X_1, X_2, \cdots, X_n)\text{为离散型} \Leftrightarrow X_i(1 \leq i \leq n)\text{为离散型}.$$

- **定义(概率质量函数)(PMF)**

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) :&= P((X_1, X_2, \cdots, X_n) = (x_1, x_2, \cdots, x_n)) \\ &= P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n). \end{aligned}$$

- **注** $\sum\limits_{(x_1, \cdots, x_n) \in \mathbb{R}} f(x_1, \cdots, x_n) \equiv 1.$

- **定义(多项分布)** 若 $B_1, B_2 \cdots, B_n$ 为互斥事件, 且 $\sum\limits_{i=1}^{n} B_i = \Omega$. 其发生的概率为 $p_1, \cdots, p_n$, 且 $\sum\limits_{i=1}^{n} p_i \equiv 1$. 满足

$$P(X_1 = k_1, \cdots, X_n = k_n) = \frac{N!}{k_1! \cdots k_n!} p_1^{k_1} \cdots p_n^{k_n},\ k_i \geq 0,\ \sum_{i=1}^{n} k_i = N.$$

其中 $\dfrac{N!}{k_1! \cdots k_n!}$ 为多项式系数.

3 连续分布

- **定义(联合 PDF)** 若存在 $f(x_1, \cdots, x_n) \geq 0$, 使得 $\forall Q \subset \mathbb{R}^n$ 可测, 都有

$$P((X_1, \cdots, X_n) \in Q) = \int_Q f(x_1, \cdots, x_n) dx_1 \cdots dx_n$$

则称 $(X_1, \cdots, X_n)$ 为连续型, f 为 $(X_1, \cdots, X_n)$ 的概率密度函数 **(PDF)**.

- **注**
  - $\int_{\mathbb{R}} f \equiv 1$;
  - 以 $n=2$ 为例, $F(a,b)=\int_{-\infty}^{a}(\int_{-\infty}^{b} f(s,t)dt)ds$;
  - $f(a,b)=\dfrac{\partial^2 F}{\partial y \partial x}(a,b)$, $a.e.$
- **定义(连续分布)(矩形域)**

$$f(x,y)=\begin{cases}\dfrac{1}{(b-a)(d-c)}, & (x,y)\in(a,b)\times(c,d)\\ 0, & otherwise\end{cases}$$

- **定义(二元正态分布)**

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

其中 $(x,y)\in\mathbb{R}^2, |\rho|<1$.

上式中 exp 的指数可视为 $-\dfrac{1}{2}\overline{X}^T W\overline{X}=-\dfrac{1}{2}\overline{AX}^T\overline{AX}$, 其 **Cholesky** 分解为

$$\overline{X}=\begin{pmatrix}\dfrac{x-\mu_1}{\sigma_1}\\ \dfrac{y-\mu_2}{\sigma_2}\end{pmatrix}, W=\frac{1}{1-\rho^2}\begin{pmatrix}1 & -\rho\\ -\rho & 1\end{pmatrix}$$

$$\Rightarrow A=\frac{1}{\sqrt{1-\rho^2}}\begin{pmatrix}\pm1 & \mp\rho\\ 0 & \pm\sqrt{1-\rho^2}\end{pmatrix}.$$

- **注**
  - $f(x,y)$ 的等值线图像为椭圆;
  - ρ 的意义?

4 边际分布

- **定义(边际 CDF)**

$$F_i(x):=P(X_i\le x)=P(X_i\le x,-\infty<X_j<\infty\,(j\ne i)).$$

- **连续型**

$n=2$ 时

$$F_X(x)=P(X\le x)=P(X\le x,-\infty<Y<\infty)=\lim_{y\to\infty}F(x,y).$$

$n=3$ 时

$$F_X(x)=P(X\le x,-\infty<Y,Z<\infty)=\lim_{y\to\infty,z\to\infty}F(x,y,z).$$

$$F_{X,Y}(x,y) = P(X \le x, Y \le y, -\infty < Z < \infty) = \lim_{z\to\infty} F(x,y,z).$$

- **离散型**

  $n = 2$ 时

$$F_X(x) = P(X \le x) = \sum_{a \le x} P(X = a) = \sum_{y} \sum_{a \le x} P(X = a, Y = y).$$

- **例(容斥原理)**

$$P(X > a, Y > b) = 1 - F_X(a) - F_Y(b) + F_{X,Y}(a,b).$$

- **定义(边际 PDF)**

$$F_X(x) := P(X \le x) = \lim_{y\to\infty} F(x,y) = \int_{-\infty}^{x} (\int_{-\infty}^{\infty} f(s,t)dt)ds$$

  $\Rightarrow X$ 的边际 **PDF** 为

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

- **例** 二元正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.
- **解答**

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, x \in \mathbb{R}.$$

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

- **注** 联合分布可确定边际分布, 边际分布不可确定联合分布.

5 条件分布 (以 $n = 2$ 为例)

- **定义(离散型条件分布)** $P(X = a_i, Y = b_j) = p_{ij} \ge 0$, $\sum\limits_{i,j} p_{ij} \equiv 1$.

$$P(X = a_i \mid Y = b_j) = \frac{P(X = a_i, Y = b_j)}{P(Y = b_j)} = \frac{p_{ij}}{\sum\limits_{k} p_{kj}}.$$

- **注** $\sum\limits_{i} P(X = a_i \mid Y = b_j) \equiv 1$.
- **定义(连续型条件分布)** (X, Y) 的 **PDF** 为 $f(x,y)$.

$$P(X \le x \mid y \le Y \le y+dy) = \frac{P(X \le x, y \le Y \le y+dy)}{P(y \le Y \le y+dy)}$$
$$= \frac{\int_{-\infty}^{x}(\int_{y}^{y+dy} f(s,t)dt)ds}{\int_{y}^{y+dy} f_Y(t)dt}.$$

- **定义(条件密度函数)**

$$f_{X|Y}(x \mid y \le Y \le y+dy) = \frac{\int_{y}^{y+dy} f(x,t)dt}{\int_{y}^{y+dy} f_Y(t)dt}.$$

令 $dy \to 0$, 定义条件密度函数:

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

条件密度函数 $f_{X|Y}(x \mid y)$ 为 **PDF**.

- **注**
  - $F(a \mid y) = P(X \le a \mid Y = y) = \int_{-\infty}^{a} f_{X|Y}(x \mid y)dx$;
  - **(乘法法则)** $f(x,y) = f_{X|Y}(x \mid y)f_Y(y)$;
  - **(全概率公式)** $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\infty}^{\infty} f_{X|Y}(x \mid y)f_Y(y)dy$;
  - **(Bayes 公式)** $f_{Y|X}(y \mid x) = \dfrac{f(x,y)}{f_X(x)} = \dfrac{f_{X|Y}(x \mid y)f_Y(y)}{\int_{-\infty}^{\infty} f_{X|Y}(x \mid y)f_Y(y)dy}$.
- **例** 二元正态分布.
- **解答** 注意到

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_2} \cdot \frac{1}{\sqrt{1-\rho^2}} \cdot \exp\{-\frac{[y - (\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))]^2}{2(1-\rho^2)\sigma_2^2}\}.$$

即当 $X = x$ 时, $Y \sim N(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), (1-\rho^2)\sigma_2^2)$.

6 独立性

- **定义(独立性)** (X,Y) 的 **CDF** 为 $F(x,y)$, 边际 **CDF** $F_X(x)$, $F_Y(y)$. 若

$$F(x,y) = F_X(x)F_Y(y), \forall\, x,y \in \mathbb{R}.$$

则称 X,Y 相互独立.

- **注** X, Y 独立 $\Leftrightarrow f(x,y) = f_X(x)f_Y(y), \forall\, x,y \in \mathbb{R}$, 其中 f 为 **PDF/PMF**.
- **定义** $X_1,\cdots,X_n$ 相互独立 $\Leftrightarrow F(x_1,\cdots,x_n) = F_1(x_1)\cdots F_n(x_n), \forall\, x_1,\cdots,x_n \in \mathbb{R}$.

- **注** $X_1, \cdots, X_n$ 独立 $\Leftrightarrow f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n), \forall x_1, \cdots, x_n \in \mathbb{R}$, 其中 f 为 **PDF/PMF**.
- **定理**
  - $f(x_1, \cdots, x_n) = g_1(x_1) \cdots g_n(x_n), \forall x_1, \cdots, x_n \in \mathbb{R}$, 则 $X_1, \cdots, X_n$ 独立.
  - $X_1, \cdots, X_n$ 独立, $Y_1 = g_1(X_1, \cdots, X_n), Y_2 = g_2(X_1, \cdots, X_n)$, 则 Y_1, Y_2 独立.

7 随机向量的函数

- **定义** $Y = g(X_1, \cdots, X_n)$
- **例** $X_i \sim B(n_i, p), i = 1, 2$ 独立, $Y = X_1 + X_2$.
- **解答**

$$\begin{aligned}
P(Y = k) = P(X_1 + X_2 = k) &= \sum_{j=0}^{k} P(X = j, X_2 = k - j) \\
&= \sum_{j=0}^{k} P(X = j) P(X_2 = k - j) \\
&= \sum_{j=0}^{k} C_{n_1}^{j} C_{n_2}^{k-j} p^k (1-p)^{n_1+n_2-k} \\
&= C_{n_1+n_2}^{k} p^k (1-p)^{n_1+n_2-k}.
\end{aligned}$$

那么有 $Y \sim B(n_1 + n_2, p)$.

- **例** X_1, X_2 连续, $X_1 > 0$, 其联合 **PDF** 为 $f(x_1, x_2)$, 且 $Y = \dfrac{X_2}{X_1}$.
- **解答** 注意到 $\forall y > 0$,

$$\begin{aligned}
P(Y \leq y) &= P(\frac{X_2}{X_1} \leq y) \\
&= P(X_2 \leq y X_1) \\
&= \int_D f(x_1, x_2) dx_1 dx_2 \\
&= \int_0^{\infty} (\int_{-\infty}^{y x_1} f(x_1, x_2) dx_2) dx_1 \\
&= \int_0^{\infty} (\int_{-\infty}^{y} f(x_1, x_1 t) x_1 dt) dx_1.
\end{aligned}$$

故 Y 的 **PDF** 为

$$l(y) = \int_0^{\infty} x_1 f(x_1, x_1 y) dx_1.$$

- **定义(密度函数变换法)** X_1, X_2 的联合 **PDF** 为 $f(x_1, x_2)$, g_1, g_2 可微可逆, 满足

$$\begin{cases} Y_1 = g_1(X_1, X_2) \\ Y_2 = g_2(X_1, X_2) \end{cases} \Rightarrow \begin{cases} X_1 = h_1(Y_1, Y_2) \\ X_2 = h_2(Y_1, Y_2) \end{cases}$$

那么

$$
\begin{aligned}
P((Y_1, Y_2) \in A) &= P((X_1, X_2) \in B) \\
&= \int_B f(x_1, x_2) dx_1 dx_2 \\
&= \int_A f(h_1(y_1, y_2), h_2(y_1, y_2))|J| dy_1 dy_2.
\end{aligned}
$$

其中

$$
J = det \begin{pmatrix} \frac{\partial h_1}{\partial y_1}, \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1}, \frac{\partial h_2}{\partial y_2} \end{pmatrix}
$$

故 Y_1, Y_2 的 **PDF** 为

$$
l(y_1, y_2) = f(h_1(y_1, y_2), h_2(y_1, y_2))|J|.
$$

- **例** X_1, X_2 连续, $X_1 > 0$, 其联合 **PDF** 为 $f(x_1, x_2)$, 且 $Y = X_1 + X_2$.
- **解答**

  令 $Z = X_1$, 则 $X_1 = Z, X_2 = Y - Z$. 故 Y, Z 的 **PDF** 为

$$
l(y, z) = f(z, y - z)|J| = f(z, y - z).
$$

  其中

$$
J = det \begin{pmatrix} 0, & 1 \\ 1, & -1 \end{pmatrix}
$$

  那么 Y 的 **PDF** 为

$$
l_Y(y) = \int_{-\infty}^{\infty} f(z, y - z) dz
$$

- **注**
  - 若 X_1, X_2 独立, 则

$$
l_Y(y) = \int_{-\infty}^{\infty} f_1(z) f_2(y - z) dz = f_1 * f_2(y).
$$

  - 若 $(X_1, X_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则

$$
Y = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2).
$$

- **注** 三大分布: **Chi-Square** 分布 $\chi^2(n)$, t_n, F.