§6 关于分布函数

对于严格递增的分布函数 F，可以考虑它的反函数 F^{-1}。现在要将其定义扩展到一般的函数，为此，对于一般的分布函数，给出如下的定义：

定义 设 $F(x)$ 是一分布函数，定义 $F^{-1}(y)$

$$F^{-1}(y) = sup\{x|F(x) < y\}, 0 < y < 1.$$

显然，$F^{-1}(y)$ 是递增的（不一定严格递增）。下面给出关于 F^{-1} 一个基本命题.

命题 $F^{-1}(y) \leqslant x \Longleftrightarrow y \leqslant F(x)$.

证 我们证明一个等价的命题

$$F^{-1}(y) > x \Longleftrightarrow y > F(x).$$

$\Longrightarrow$：由 $F^{-1}(y) > x$ 知 $\exists x^*0$ 使得 $x < x^* < F^{-1}(y)$。故由分布函数的单调性可知

$$F(x) \leqslant F(x^*) < y.$$

$\Longleftarrow$：由 $y > F(x)$ 及分布函数的右连续性知 $\exists \varepsilon > 0$ 使得 $y > F(x+\varepsilon)$，于是有

$$F^{-1}(y) \geqslant x + \varepsilon > x.$$

然后就可以得到下面的这条很基本的定理：

定理 设 $F(x)$ 是分布函数，$X \sim U[0,1]$，则 $Y = F^{-1}(X)$ 的分布函数是 $F(\cdot)$。

证

$$\begin{aligned} P(Y \leqslant y) &= P(F^{-1}(X) \leqslant y) \\ &= P(X \leqslant F(y)) \\ &= F(y). \end{aligned}$$

$F^{-1}(\cdot)$ 还有以下这些性质：

(1) $F(F^{-1}(y))) \geqslant y$;

(2) $F^{-1}(F(y)) \leqslant y$;

(3) 若 $F(x)$ 在 $x = F^{-1}(y)$ 处连续，则 $F(F^{-1}(y)) = y$;

(4) $F^{-1}(y)$ 左连续。

第三章 随机向量及其概率分布

§1 连续型随机向量及其概率密度函数

定义 称概率空间 $(\Omega, \mathcal{F}, P)$ 上 n 个随机变量 $X_1, X_2, \cdots, X_n$ 的整体 $X = (X_1, X_2, \cdots, X_n)$ 为 n 维随机向量。

很多随机现象中往往涉及到多个随机变量，例如炮弹的落点 (X, Y) 就是两个随机变量组成的一个随机向量。随机向量可以是有限维的，也可以是无穷维的。

一个 n 的随机向量 $X = (X_1, X_2, \cdots, X_n)$ 可以看作如下的映射：

$$\begin{aligned} \Omega &\longrightarrow \mathbb{R}^n; \\ \omega &\longrightarrow (X_1(\omega), X_2(\omega), \cdots, X_n(\omega)). \end{aligned}$$

其中，每一个分量都是可测的，即

$$\{\omega | X_i(\omega) \leqslant a_i\} \in \mathcal{F}.$$

因此，

$$\begin{aligned} & \bigcap_{i=1}^{n} \{\omega | X_i(\omega) \leqslant a_i\} \\ = \ & \{\omega | X_1(\omega \leqslant a_i, X_2(\omega) \leqslant a_2, \cdots, X_n(\omega) \leqslant a_n\} \in \mathcal{F}. \end{aligned}$$

于是，$P(X_1 \leqslant a_1, X_2 \leqslant a_2, \cdots, X_n \leqslant a_n)$ 就有意义了。

同样的，定义 n 维联合分布函数为

$$F(a_1, a_2, \cdots, a_n) = P(X_1 \leqslant a_1, X_2 \leqslant a_2, \cdots, X_n \leqslant a_n).$$

同一维的分布函数一样，联合分布函数有以下性质：

(1) $\lim\limits_{a_i \to -\infty} F(a_1, a_2, \cdots, a_n) = 0, \quad i = 1, 2, \cdots, n;$

$\lim\limits_{\substack{a_1 \to +\infty \\ a_2 \to +\infty \\ \cdots \\ a_n \to +\infty}} F(a_1, a_2, \cdots, a_n) = 1;$

(2) F 对每个分量都是单调非降的；

(3) F 是右连续的，且左极限存在。

下面定义连续型的随机向量。

定义 称 n 维随机向量 $(X_1, X_2, \cdots, X_n)$ 是连续型的，如果存在非负可积函数 $p(x_1, x_2, \cdots, x_n)$，使得对任意的 $x_1, x_2, \cdots, c_n$，都有

$$F(x_1, x_2, \cdots, x_n) = \int \cdots \int_D p(t_1, t_2, \cdots, t_n) \mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

其中，$D = \{(t_1, \cdots, t_n) | t_1 \leqslant x_1, \cdots, t_n \leqslant x_n\}$，并称 $p(x_1, \cdots, x_n)$ 为 $X_1, \cdots, X_n$ 的联合密度函数。

容易看出，联合密度函数并不唯一，改变可数个点的函数值还是密度函数。

下面给出两个最基本的分布。

(1) 均匀分布

称 $(X_1, X_2, \cdots, X_n)$ 为 n 维区域 G 上的均匀分布，如果其联合密度为

$$p(x_1, \cdots, x_n) = \begin{cases} c, & (x_1, \cdots, x_n) \in G; \\ 0, & \text{其他}. \end{cases}$$

可以看出，均与分布实际上就是几何概型。

(2) 多维正态分布

先定义二维的。

如果 (X, Y) 有联合密度

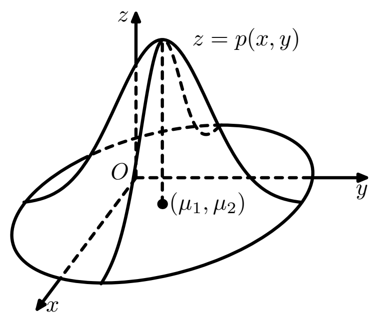
$$\begin{aligned} p(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\Big\{ -\frac{1}{2(1-\rho^2)} \\ &\quad \cdot \Big[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\Big]\Big\}. \end{aligned}$$

其中，$-\infty < \mu_1 < \infty$，$-\infty < \mu_2 < \infty$，$\sigma_1 > 0$，$\sigma_2 > 0$，$-1 < \rho < 1$，则称 (X, Y) 服从二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$。

然后再定义高维的。

称 n 维随机向量 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布，如果它具有密度

$$p(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \cdot \exp\Big[-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\Big],$$



其中，$x=(x_1,x_2,\cdots,)'$，而 $\mu=(\mu_1,\mu_2,\cdots,\mu_n)'$ 是参数向量，$\Sigma=(\sigma_{ij})_{n\times n}$ 是 n 阶正定矩阵。用记号 $X\sim N(\mu,\Sigma)$。

容易看出，取

$$\mu=(\mu_1,\mu_2)' \quad , \quad \Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix},$$

就是二维正态分布。

§2 离散型随机向量及边缘分布函数

定义 称 n 维随机向量 $X=(X_1,X_2,\cdots X_n$ 是离散型的，如果它只取至多可数个不同的值。

相应的，以二维为例，其概率分布列定义如下：

$$p_{ij}=P((X_1,X_2)=(x_{1i},x_{2j})),\quad i=1,2,\cdots,\quad j=1,2,\cdots.$$

其中 p_{ij} 满足

(1) $p_{ij}\geqslant 0,\ i=1,2,\cdots,\ j=1,2,\cdots;$

(2)$\sum\limits_{i,j}p_{ij}=1.$

离散型随机向量中，常见的是多项分布，其定义如下：

称 $(X_1,X_2,\cdots,X_r)$ 服从参数为 $(n;p_1,p_2,\cdots,p_r)$ 的多项分布，如果

$$P((X_1,\cdots,X_r)=(k_1,\cdots,k_r))=\frac{n!}{k_1!\cdots k_r!}p_1^{k_1}\cdots p_r^{k_r},$$

其中 $k_1,\cdots,k_n$ 为非负整数且 $\sum\limits_{i=1}^{r}k_i=n$，$\sum\limits_{i=1}^{r}p_i=1$。

例 2.1 盒中有 2 个红球，3 个白球，5 个黑球共 10 个球，从中有放回的取 7 次，若记 X_1, X_2, X_3 分别为取到的红球，白球和黑球的数目，则 (X_1, X_2, X_3) 是一离散型随机向量，且服从参数为 $(7; 0.2, 0.3, 0.5)$ 的多项分布。

定义 对于一个多维随机向量 $X = (X_1, X_2, \cdots, X_n)$，它的一些分量 $(X_{i_1}, \cdots, X_{i_k}), 1 \leqslant i_1 < i_2 < \cdots < i_k, 1 \leqslant k < n$ 组成一个新的随机向量，这个新的随机向量的分布函数就称为边缘分布函数。

对于离散型随机变量，以二维为例，设 (X, Y) 的联合分布为

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

则 X 的边缘分布为

$$P(X = x_i) = \sum_j P((X,Y) = (x_i, y_j)) = \sum_j p_{ij}, \quad i = 1, 2, \cdots.$$

同样，Y 的边缘分布为

$$P(Y = y_j) = \sum_i p_{ij}, \quad j = 1, 2, \cdots.$$

对于连续型随机变量，它的边缘分布函数也有密度函数。以二维为例，设 (X, Y) 的联合密度函数为 $p(x, y)$，则边缘密度函数为

$$p_X(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y, \quad p_Y(y) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}x.$$

证 直接根据分布函数的定义，

$$\begin{aligned} F_X(t) &= P(X < t) \\ &= P(X < t, -\infty < y < \infty) \\ &= \iint_{(-\infty,t)\times(-\infty,\infty)} p(x,y)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{t} \Big[\int_{-\infty}^{\infty} p(x,y)\mathrm{d}y\Big]\mathrm{d}x. \end{aligned}$$

根据密度函数的定义可知 $p_X(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y$。同理 $p_Y(y) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}x$。

下面看一个例子。

例 2.1 设 $(X,y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求 X,Y 的边缘密度。

解 直接代入密度函数

$$
\begin{aligned}
p_X(x) &= \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\big\{-\frac{1}{2(1-\rho^2)}\frac{(x-\mu_1)^2}{\sigma^2}\big\} \\
&\quad\cdot\int_{-\infty}^{\infty}\exp\big\{-\frac{1}{2(1-\rho^2)}\big[-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\big]\big\}\mathrm{d}y
\end{aligned}
$$

对积分式中部分进行配方整理，得

$$
\begin{aligned}
p_X(x) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\big\{-\frac{1}{2(1-\rho^2)}\frac{(x-\mu_1)^2}{\sigma^2}+\frac{\rho^2(x-\mu_1)^2}{2(1-\rho^2)\sigma_1^2}\big\} \\
&\quad\cdot\int_{-\infty}^{\infty}\exp\big\{-\frac{1}{2(1-\rho^2)}\big[\frac{y-\mu_2}{\sigma_2}-\frac{\rho(x-\mu_1)}{\sigma_1}\big]^2\big\}\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\cdot\exp\big\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\big\}\cdot\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \\
&\quad\cdot\int_{-\infty}^{\infty}\exp\big\{-\frac{1}{2(1-\rho^2)\sigma_2^2}\big[y-\mu_2-\frac{\rho\sigma_2(x-\mu_1)}{\sigma_1}\big]^2\big\}\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\big\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\big\}.
\end{aligned}
$$

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

知道了边缘分布，一个很直接的问题是，能不能确定联合分布？答案是否定的。下面给出一个离散情况的例子。

例 2.2 一个密封的盒子中放有 2 个白球，3 个黑球。记

$$
\xi=\begin{cases}1, & \text{第一次摸到白球,}\\ 0, & \text{第一次摸到黑球;}\end{cases}
$$

$$
\eta=\begin{cases}1, & \text{第二次摸到白球,}\\ 0, & \text{第二次摸到黑球,}\end{cases}
$$

计算可知，在有放回和无放回两种情形下，ξ 和 η 都服从参数为 0.4 的二项分布，但有放回和无放回时 (ξ,η) 的联合分布是不同的。