§3 独立性

在第一章里面讨论了事件的相互独立性，现在要在这个基础上，进一步讨论随机变量的相互独立性。

定义 设 $X_1, \cdots, X_n$ 是概率空间 $(\Omega, \mathcal{F}, P)$ 上的 n 个随机变量，称它们是相互独立的如果对任意 n 个实数 $x_1, \cdots, x_n$ 都有

$$P(X_1 \leqslant x_1, \cdots, X_n \leqslant x_n) = P(X_1) \cdots P(X_n).$$

实际上，随机变量的独立性也是事件的独立性，上式等价于事件 $\{X_1 \leqslant x_1\}, \cdots, \{X_n \leqslant x_n\}$ 相互独立。

关于随机变量的相互独立有下面的性质：

如果 $X_1, \cdots, X_n$ 相互独立，$f_1(x), \cdots, f_n(x)$ 都是一元 $Borel$ 可测函数，则 $f_1(X_1), \cdots, f_n(X_n)$ 也是相互独立的。

对于连续型随机变量 $X_1, \cdots, X_n$，设其密度函数分别 $p_1(x_1), \cdots, p_n(x_n)$，则

$$\begin{aligned} & X_1, \cdots, X_n \text{相互独立} \\ \Longleftrightarrow & \prod_{i=1}^{n} p_i(x_i) \text{是} (X_1, \cdots, X_n) \text{的联合密度}. \end{aligned}$$

上一节给出了二元正态分布的边缘分布，现在来看它的两个分量何时独立。

例 3.1 设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$，则 X, Y 相互独立的充要条件是 $\rho = 0$。

证 X, Y 的边缘密度函数分别为

$$P_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \cdot \exp\Big\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\Big\}, \quad P_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \cdot \exp\Big\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\Big\}.$$

当 $\rho = 0$ 时，有

$$p(x, y) = p_X(x) p_Y(y).$$

可知 X, Y 相互独立，充分性得证。

若 X, Y 相互独立，由于 $p(x,y), p_X(x), p_Y(y)$ 都是连续函数，故对任意的 (x,y) 都有

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

特别的取 $(X,Y)=(\mu_1,\mu_2)$，有

$$\frac{1}{2\pi\sigma_1\sigma_2} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \quad .$$

由此推出 $\rho = 0$。必要性得证。

这里用到了密度函数一个性质：如果 $p_1(x,y), p_2(x,y)$ 都是 (X,Y) 的密度函数，且它们都在 (x_0,y_0) 处连续，则必有

$$p_1(x_0,y_0) = p_2(x_0,y_0).$$

§4 随机向量函数及其分布

设 n 维随机向量 $X=(X_1,\cdots,X_n)$ 的联合分布函数为 $F_X(x_1,\cdots,x_n)$，而

$$\begin{aligned}
&Y_1 = g_1(X_1,\cdots,X_n),\\
&Y_2 = g_2(X_1,\cdots,X_n),\\
&\cdots\\
&Y_m = g_m(X_1,\cdots,X_n),
\end{aligned}$$

其中 $g_i(x_1,\cdots,x_n), i=1,\cdots,m$ 都是 n 元 $Borel$ 可测函数，现在要求出 m 维随机向量 $Y=(Y_1,\cdots,Y_m)$ 的联合分布函数 $F_Y(y_1,\cdots,y_m)$。

一般的解法如下：记 $C=\{(x_1,\cdots,x_n)|g_i(x_1,\cdots,x_n)\leqslant y_i,\ i=1,\cdots,\}$，

则

$$\begin{aligned}
F_Y(y_1,\cdots,y_m) &= P(g_1(X_1,\cdots,X_n)\leqslant y_1,\cdots,g_m(X_1,\cdots,X_n)\leqslant y_m)\\
&= P((X_1,\cdots,X_n)\in C).
\end{aligned}$$

当 X 有联合密度 $p(x_1,\cdots,x_n)$ 时，有

$$F_Y(y_1,\cdots,y_m) = \int\cdots\int_C p(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

下面给出几个例子。

例 4.1 设 (X,Y) 的联合密度为 $p(x,y)$，求 $Z=X+Y$ 的分布。

解 直接计算 Z 的分布函数

$$\begin{aligned}F_Z(z) &= P(X+Y\leqslant z)\\ &= \iint\limits_{x+y\leqslant z} p(x,y)\mathrm{d}x\mathrm{d}y\\ &= \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{z-x}p(x,y)\mathrm{d}y\\ &= \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{z}p(x,u-x)\mathrm{d}u\\ &= \int_{-\infty}^{z}\mathrm{d}u\int_{-\infty}^{\infty}p(x,u-x)\mathrm{d}x\end{aligned}$$

这表明 Z 具有密度函数

$$p_Z(z)=\int_{-\infty}^{\infty}p(x,z-x)\mathrm{d}x.$$

由对称性可知，$p_Z(z)=\int_{-\infty}^{\infty}p(z-y,y)\mathrm{d}y$。

例 4.2 设 X,Y 相互独立，且都服从分布 $N(\mu,\sigma^2)$，求 $X+Y$ 的分布密度函数 $g(z)$。

解 由上例直接计算可得

$$g(z)=\frac{1}{\sqrt{2\pi}\sqrt{2\sigma}}\cdot\exp\big\{-\frac{(z-2\mu)^2}{4\sigma^2}\big\}.$$

这表明 $X+Y\sim N(2\mu,2\sigma^2)$。

例 4.3 设 X,Y 相互独立且都服从分布 $N(0,1)$，求 $Z=\sqrt{X^2+Y^2}$ 的分布密度。

解 用极坐标变化

$$\begin{cases}x=r\cos\theta,\\ y=r\sin\theta,\end{cases}\qquad r\geqslant 0, 0\leqslant\theta\leqslant 2\theta.$$

$$
\begin{aligned}
F_Z(z) &= P(\sqrt{X^2+Y^2}\leqslant z)\\
&= \iint\limits_{\sqrt{X^2+Y^2}\leqslant z} \frac{1}{2\pi}\cdot\exp\{-\frac{1}{2}(x^2+y^2)\}\mathrm{d}x\mathrm{d}y.\\
&= \int_0^{2\pi}\int_0^z \frac{1}{2\pi}\mathrm{e}^{-\frac{r^2}{2}} r\mathrm{d}r\\
&= \int_0^z r\mathrm{e}^{-\frac{r^2}{2}}\mathrm{d}r
\end{aligned}
$$

因此，

$$
p_Z(z)=\begin{cases} z\mathrm{e}^{-\frac{z^2}{2}}, & z\geqslant 0;\\ 0, & \text{其它}. \end{cases}
$$

这个分布称为 *Rayleigh* 分布。

上面的都是 $2\to 1$ 的情形，下面要讨论 $n\to n$ 的情形，以 $n=2$ 为例。

设 (X,Y) 有联合密度 $p(x,y)$，且区域 A（可以使全平面）满足 $P((X,Y)\in A)=1$；对变换 (Δ)；

$$
\begin{cases} u=f(x,y),\\ v=g(x,y), \end{cases}
$$

当 $(x,y)\in A$ 时，(u,v) 的值域为 G，且 (Δ) 满足：

(1) $A\xrightarrow{(\Delta)} G$ 是一一对应的；

(2) f,g 在 A 中有连续偏导数；

(3) $\dfrac{\partial(u,v)}{\partial(x,y)}$ 在 A 中处处不为 0；

则 $U(=f(X,Y)),V(=g(X,Y))$ 具有联合密度函数：

$$
q(u,v)=\begin{cases} p(x(u,v),y(u,v))\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right|, & (u,v)\in G;\\ 0, & \text{其它}. \end{cases}
$$

其中，$x(u,v),y(u,v)$ 是由 (Δ) 决定的反函数，而 $\dfrac{\partial(u,v)}{\partial(x,y)}$ 为相应的雅可比行列式。

例 4.4 设 X,Y 相互独立，且都服从 $N(0,1)$，(R,Θ) 是平面上随机点 (X,Y) 相应的极径、极角，即有

$$
\begin{cases} X=R\cos\Theta,\\ Y=R\sin\Theta, \end{cases} \quad R\geqslant 0, 0\leqslant\Theta\leqslant 2\pi.
$$

求 (R,Θ) 的联合密度。

解 (X,Y) 到 (R,Θ) 的变换 (Δ) 为

$$\begin{cases} R=\sqrt{X^2+Y^2}, \\ \Theta=\arctan\dfrac{Y}{X}, \end{cases}$$

相应的雅可比行列式

$$\left|\frac{\partial(u,v)}{\partial(x,y)}\right|=\frac{1}{r}.$$

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=r.$$

于是，(R,Θ) 的联合密度函数为

$$\begin{aligned} q(r,\theta) &= \frac{1}{2\pi}\cdot \mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\left|\frac{\partial(x,y)}{\partial(u,v)}\right| \\ &= \frac{1}{2\pi}\cdot \mathrm{e}^{-\frac{1}{2}r^2}\cdot r, \quad r>0, \quad 0<\theta<2\pi. \end{aligned}$$

还可以看出，R 和 Θ 是相互独立的，R 服从 $Rayleigh$ 分布，而 Θ 服从 $(0,2\pi)$ 上的均匀分布。

这个例子的结果还可以用来产生独立的两个正态随机数，具体方法如下：先产生两个相互独立的 $(0,1)$ 的均匀分布的随机数 U_1,U_2，然后取

$$X=(-2\ln U_1)^{\frac{1}{2}}\cos 2\pi U_2,$$

$$Y=(-2\ln U_1)^{\frac{1}{2}}\sin 2\pi U_2,$$

则 X,Y 是相互独立的 $N(0,1)$ 随机数。

接下来高维正态分布的线性函数的分布。

定理 设 $\mathbf{X}\sim N(\mu,\Sigma)$，$B$ 是任一 n 阶非退化矩阵，则

$$\mathbf{Y}=B\mathbf{X}\sim N(B\mu,B\Sigma B').$$

证明直接计算 $\mathbf{Y}$ 的分布函数可得。

再给出一个关于高维正态分布的边缘分布函数的定理。

定理 设 $\mathbf{X}\sim N(\mu,\Sigma)$，若

$$\Sigma=\begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix},$$

则

$$\mathbf{X_1} \sim N(\mu_{\mathbf{1}}, \Sigma_1), \quad \mathbf{X_2} \sim N(\mu_{\mathbf{2}}, \Sigma_2).$$

其中，Σ_1 是 m 阶方阵，Σ_2 是 $n-m$ 阶方阵，而 $\mathbf{X_1}, \mathbf{X_2}, \mu_{\mathbf{1}}, \mu_{\mathbf{2}}$ 是相应的列向量。

由上面的两个定理就可以退出一般情形的高维正态分布的边缘分布了。

定理 设 $\mathbf{X} \sim N(\mu, \Sigma)$，对 Σ 进行分块

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

则

$$\mathbf{X_1} \sim N(\mu_{\mathbf{1}}, \Sigma_1), \quad \mathbf{X_2} \sim N(\mu_{\mathbf{2}}, \Sigma_2).$$

其中，Σ_1 是 m 阶方阵，Σ_2 是 $n-m$ 阶方阵，而 $\mathbf{X_1}, \mathbf{X_2}, \mu_{\mathbf{1}}, \mu_{\mathbf{2}}$ 是相应的列向量。利用矩阵的变换就可以证明。

对于 n 个独立同分布的服从 $N(0,1)$ 的随机变量，它们的平方和服从的分布在统计里面应用很广。下面来计算这个分布。

设 $X_1, \cdots, X_n$ 相互独立且同分布，服从 $N(0,1)$，求 $Y = X_1^2 + \cdots + X_n^2$ 的分布密度函数。

解 记 $D_y = \{(x_1. \cdots, x_n) | x_1^2 + \cdots + x_n^2 \leqslant y\}$，则

$$\begin{aligned} P(Y \leqslant y) &= P(X_1^2 + \cdots X_n^2 \leqslant y) \\ &= \int \cdots \int_{D_y} \frac{1}{(\sqrt{2\pi})^n} \cdot \exp\{ -\frac{x_1^2 + \cdots + x_n^2}{2} \} \mathrm{d}x_1 \cdots \mathrm{d}x_n \end{aligned}$$

进行 n 维球坐标变换

$$\begin{cases} x_1 = r\cos\theta_1, \\ x_2 = r\sin\theta_1\cos\theta_2, \\ \cdots \\ x_{n-1} = r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\cos\theta_{n-1}, \\ x_n = r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\sin\theta_{n-1}, \end{cases}$$

其中，$r > 0, 0 \leqslant \theta_1, \theta_2, \cdots, \theta_{n-2} < \pi, 0 \leqslant \theta_{n-1} < 2\pi$。于是上面的积分区域就化为

$$D_y^* = \{(r, \theta_1, \cdots, \theta_{n-1}) | 0 < r \leqslant \sqrt{y}, 0 \leqslant \theta_1, \theta_2, \cdots, \theta_{n-2} < \pi, 0 \leqslant \theta_{n-1} < 2\pi\}.$$

于是上面的积分式变为

$$\begin{aligned} P(Y \leqslant y) &= \int \cdots \int_{D_y^*} \frac{1}{(\sqrt{2\pi})^n} \cdot \mathrm{e}^{-\frac{r^2}{2}} r^{n-1} f(\theta_1, \cdots, \theta_{n-1}) \mathrm{d}r \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_{n-1} \\ &= c \cdot \int_0^{\sqrt{y}} r^{n-1} \mathrm{e}^{-\frac{r^2}{2}} \mathrm{d}r \\ &= c \cdot \int_0^y s^{\frac{n}{2}-1} \mathrm{e}^{-\frac{s}{2}} \mathrm{d}s. \end{aligned}$$

于是，Y 有密度函数

$$p_Y(y) = c y^{\frac{n}{2}-1} \mathrm{e}^{-\frac{y}{2}},$$

其中 c 为归一化常数。比较以上密度与 $\Gamma(\frac{n}{2}, \frac{1}{2})$ 的密度函数，可知 $Y \sim \Gamma(\frac{n}{2}, \frac{1}{2})$。

上面得到的分布称为自由度为 n 的卡方分布，记作 $\chi^2(n)$。

然后，给出一个有关卡方分布的定理。

定理 设 $X_1, \cdots, X_n$ 独立同分布 $N(0,1)$，则

(1) $\bar{X} = \dfrac{1}{n}(X_1 + \cdots + X_n) \sim N(0,1)$;

(2) $\displaystyle\sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1)$;

(3) $\bar{X}$ 与 $\displaystyle\sum_{i=1}^n (X_i - \bar{X})^2$ 相互独立。

§5 顺序统计量

定义 设 $X_1, \cdots, X_n$ 独立同分布 $F(\cdot)$, 对每个 $\omega \in \Omega$, 将 $X_1(\omega), \cdots, X_n(\omega)$ 从小到大排序：

$$X_{(1)}(\omega) \leqslant X_{(2)}(\omega) \leqslant \cdots X_{(n)}(\omega),$$

这样我们得到 n 个新的随机变量 $X_{(1)}(\omega), \cdots, X_{(n)}(\omega)$, 称它们为 $X_1, \cdots, X_n$ 的顺序统计量。

接下来的问题就是，顺序统计量的分布是什么。对于 $X_{(1)}, X_{(n)}$，分布比较容易求。利用

$$X_{(1)} = \min X_1, \cdots, X_n,$$

$$X_{(n)} = \max X_1, \cdots, X_n,$$

可知

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_{(n)} \leqslant x) \\ &= P(X_1 \leqslant x, \cdots, X_n \leqslant x) \\ &= [F(x)]^n. \end{aligned}$$

类似的可求出 $X_{(1)}$ 的分布函数。

对于一般的 k，也可以跟上面的类似的计算分布函数，只是分析要复杂一些，可求得

$$F_{X_{(k)}}(x) = \sum_{i=k}^{N} C_n^i (F(x))^i (1-F(x))^{n-i}, \quad k = 1, 2, \cdots, n.$$

用分步积分公式可以验证下面的恒等式：

$$\sum ni = kC_i^n p^i (1-p)^{n-i} = \frac{n!}{(k-1)!(n-k)!} \int_0^p t^{k-1}(1-t)^{n-k} \mathrm{d}t,$$

因此，$X_{(k)}$ 的分布函数可以改写为

$$F_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} \int_0^{F(x)} t^{k-1}(1-t)^{n-k} \mathrm{d}t.$$

当 $F(x)$ 有密度函数 $p(x)$ 时，$F_{X_{(k)}}$ 也有相应的密度函数：

$$p_{(k)}(x) = \frac{n!}{(k-1)!(n-k)!}(F(x))^{k-1}(1-F(x))^{n-k} \cdot p(x).$$

然后再看 $X_{(1)}, \cdots, X_{(n)}$ 的联合分布。当 $F(x)$ 有密度函数 $p(x)$ 时，用微元密度法可以求得 $X_{(1)}, \cdots, X_{(n)}$ 的联合密度为

$$g(x_1, \cdots, x_n) = \begin{cases} n!p(x_1)\cdots p(x_n), & x_1 < \cdots < x_n; \\ 0, & \text{其它}. \end{cases}$$