- 回顾：两个事件独立

设 A, B 是随机试验中的两个事件，若满足 $P(AB) = P(A)P(B)$，则称事件 A 和 B 相互独立.

- 对二维随机变量(X,Y)，联合分布也是两个事件同时发生的概率

$$F(x,y) = \mathbb{P}(\{X \le x\} \cap \{Y \le y\})$$

定义 3.9 相互独立的二维随机变量

设随机变量 X, Y 的联合分布为 $F(x,y)$，边缘分布为 $F_1(x), F_2(y)$. 若对 $\forall (x,y) \in \mathbb{R}^2$，都有

$$F(x,y) = F_1(x)F_2(y), \tag{3.15}$$

则称随机变量 X, Y 相互独立. ♣

联合分布函数为每个随机变量边缘分布函数的乘积

- 若(X,Y)为二维离散型随机变量，分布律为

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

则X,Y相互独立等价于 $p_{ij}=p_{i\cdot}p_{\cdot j} \quad \forall i,j=1,2,\cdots.$

- 若(X,Y)为二维连续型随机变量，有联合概率密度函数$f(x,y)$和边缘概率密度函数$f_1(x), f_2(y)$, 则X,Y相互独立等价于 $f(x,y)=f_1(x)f_2(y) \quad \forall (x,y)\in\mathbb{R}^2.$

- 随机变量独立的定义也等价于：对 $\forall B_1\in\mathbb{R}, B_2\in\mathbb{R}$, 都有

$$\mathbb{P}(X\in B_1, Y\in B_2)=\mathbb{P}(X\in B_1)\mathbb{P}(Y\in B_2).$$

- 若X,Y相互独立，它们的函数$g_1(X), g_2(Y)$也独立.

若 A 是与随机变量 X 相关的任意事件,B 是与随机变量 Y 相关的任意事件，则事件 A,B 相互独立.

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

设 (X,Y) 服从矩形 $D=[a,b]\times[c,d]$ 上的均匀分布，则 X 与 Y 相互独立。

设 (X,Y) 服从单位圆上的均匀分布，则 X 与 Y 不独立。

推广至n个随机变量相互独立

定义 3.10 多维随机向量的相互独立性

设 n 维随机变量 $(X_1, X_2, \ldots, X_n) \sim F(x_1, x_2, \ldots, x_n)$, 记 $X_i \sim F_i(x_i)$, $i = 1, 2, \ldots, n$. 若对任意 $(x_1, x_2, \ldots, x_n)' \in \mathbb{R}^n$ 都成立

$$F(x_1, x_2, \ldots, x_n) = F_1(x_1)F_2(x_2)\cdots F_n(x_n),$$

则称随机变量 $X_1, X_2, \ldots, X_n$ 相互独立. ♣

性质： 1. 若随机向量 $(X_1, X_2, \ldots, X_n)$ 是 n 维离散型随机向量，如果对所有可能取值 $(x_{1k_1}, x_{2k_2}, \cdots, x_{nk_n})$ 都成立

$$\mathbb{P}(X_1 = x_{1k_1}, X_2 = x_{2k_2}, \ldots, X_n = x_{nk_n}) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_{ik_i}),$$

则称随机变量 $X_1, X_2, \ldots, X_n$ 相互独立.

2. 设连续型随机向量 $(X_1, X_2, \ldots, X_n) \sim f(x_1, x_2, \ldots, x_n)$, $X_i \sim f_i(x_i)$, $i = 1, 2, \ldots, n$, 若对任意 $(x_1, x_2, \ldots, x_n)' \in \mathbb{R}^n$ 都成立

$$f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i),$$

则称随机变量 $X_1, X_2, \ldots, X_n$ 相互独立.

3. 若 n 个随机变量 $X_1, X_2, \ldots, X_n$ 相互独立，则随机向量 $(X_1, X_2, \ldots, X_k)$ 和随机向量 $(X_{k+1}, \ldots, X_n)$ 相互独立. 当然随机向量的函数 $Y_1 = g_1(X_1, X_2, \ldots, X_k)$ 和 $Y_2 = g_2(X_{k+1}, \ldots, X_n)$ 也是相互独立的.

- 但一般来说，仅由某一部分独立无法推出n个随机变量相互独立.

例子：若 X,Y 相互独立，都服从-1 和 1 这两点上的等可能分布，而 $Z = XY$. 则Z,X,Y 两两独立但不相互独立.

- 以二维连续型随机向量的函数为例.设$(X,Y)\sim f(x,y)$为二维连续型随机变量, 研究$Z=g(X,Y)$的分布, 函数g: 从二维到一维的变换或从二维到二维的变换.

- $Z=g(X,Y)$ 为一维随机变量, Z 的分布函数 F_Z 为

$$F_Z(z)=\mathbb{P}(Z\le z)=\iint_{g(x,y)\le z} f(x,y)dxdy.$$

求导得密度函数.

- $\mathbf{Z}=(Z_1,Z_2)=(g_1(X,Y),g_2(X,Y))$ 为二维随机向量, (Z_1,Z_2) 的联合分布函数 $F_{\mathbf{Z}}(z_1,z_2)$ 为

$$F_{\mathbf{Z}}(z_1,z_2)=\mathbb{P}(Z_1\le z_1,Z_2\le z_2)=\iint_{\substack{g_1(x,y)\le z_1\\ g_2(x,y)\le z_2}} f(x,y)dxdy$$

求导得密度函数.

- **(二维密度变换公式)** 设 (X,Y) 是二维连续型随机向量，具有联合密度 $f(x,y)$. $(Z_1,Z_2)=(g_1(X,Y),g_2(X,Y))$. 若 (Z_1,Z_2) 和 (X,Y) 一一对应，反函数记为 $X=h_1(Z_1,Z_2),Y=h_2(Z_1,Z_2)$ 且均有一阶连续偏导数. 则 (Z_1,Z_2) 亦为二维连续型随机向量，且其联合概率密度为

$$f_{\mathbf{Z}}(z_1,z_2)=f\big(h_1(z_1,z_2),h_2(z_1,z_2)\big)|J|,\quad (z_1,z_2)\in\mathbb{D}$$

其中$\mathbb{D}$是随机向量 $\mathbf{Z}=(Z_1,Z_2)$ 的密度非零的所有可能值的集合, J 是变换的Jaccobi(雅可比)行列式,即

$$J=\begin{vmatrix}\frac{\partial h_1}{\partial z_1} & \frac{\partial h_1}{\partial z_2}\\ \frac{\partial h_2}{\partial z_1} & \frac{\partial h_2}{\partial z_2}\end{vmatrix}$$

在多元随机变量场合, 更一般地有如果 $\mathbf{X}=(X_1,\cdots,X_n)$ 是 n 维连续型随机向量, 具有联合概率密度函数 $f(\mathbf{x})=f(x_1,\cdots,x_n)$. 假设存在 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的一一映射 $\mathbf{g}$, 其逆映射 $\mathbf{g}^{-1}$ 存在一阶连续偏导数, 那么 n 维随机变量 $Y=\mathbf{g}(\mathbf{X})$ 是连续型的, 且具有联合密度函数

$$p(\mathbf{y})=f\left(\mathbf{g}^{-1}(\mathbf{y})\right)|\mathbf{J}|I_{\mathbb{D}}(\mathbf{y}),$$

其中 $\mathbb{D}\subset\mathbb{R}^n$ 是随机向量 $\mathbf{y}$ 的密度非零的所有可能值的集合, J 是变换的 $Jaccobi$ 行列式, 即

$$\mathbf{J}=\left|\frac{\partial\mathbf{g}^{-1}}{\partial(y_1,\ldots,y_n)}\right|.$$

三类函数:

(1). 随机变量的和

(2). 随机变量的商

(3). 随机变量的最值

(1). 两个随机变量的和

- 设X,Y的联合概率密度为$f(x,y)$，则$Z=X+Y$的概率密度$p(z)$为

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

- 特别，当 X 与 Y 独立时，分别记 X 和 Y 的概率密度为 $f_1(x)$ 和 $f_2(y)$，则 $X+Y$ 的概率密度为

$$p(z)=\int_{-\infty}^{\infty} f_1(x)f_2(z-x)dx=\int_{-\infty}^{\infty} f_1(z-y)f_2(y)dy \quad \triangleq \quad f_1 * f_2(z)=f_2 * f_1(z)$$

称此公式为**卷积公式**。

- 当 ξ,η 是相互独立的非负整值随机变量，各有分布律 $\{a_k\}$ 与 $\{b_k\}$. 那么 $\xi+\eta$ 有分布律

$$P(\xi+\eta=n)=\sum_{k=0}^{n} a_k b_{n-k}$$

称此公式为**离散卷积公式**

- 设X服从指数分布$Exp(1/2), Y{\sim}U(0,1)$, 且X和Y相互独立。

  求$(X-Y)$的概率密度函数和$P(X \leq Y)$。

设 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$ 且 X 与 Y 相互独立，则:

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

更一般地，设 $X_i \sim N(\mu_i, \sigma_i^2), i = 1, \cdots, n$, $X_1, \cdots, X_n$ 相互独立. $a_1, \cdots, a_n, b$ 为任意 $n+1$ 个实数，其中 $a_1, \cdots, a_n$ 不全为零. 令 $X = \sum_{i=1}^{n} a_i X_i + b$，则有: $X \sim N(\mu, \sigma^2)$，其中 $\mu = \sum_{i=1}^{n} a_i \mu_i + b$, $\sigma^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$.

- 如果相互独立的两个同类型随机变量之和仍服从同一类型的分布, 则称此分布类型具有**再生性**.
- 具有再生性的一些分布:
  - 二项分布（关于试验次数具有再生性）
  - Poisson分布（关于参数λ具有再生性）
  - 负二项分布（关于成功次数r具有再生性）
  - 正态分布（关于两个参数都具有再生性）
  - χ_n^2分布（关于自由度n具有再生性）
  - Γ分布（关于参数n具有再生性）

定义 5.6 (卡方分布)

设样本 $(X_1, X_2, \ldots, X_n)$ 为标准正态总体的一个简单随机样本, 称

$$X = X_1^2 + X_2^2 + \cdots + X_n^2$$

是自由度为 n 的 χ^2 分布, 记为 $X \sim \chi_n^2$. ♣

如果 $X_1, \ldots, X_n$ 相互独立且服从相同的指数分布 $Exp(\lambda)$, 则

$$S_n = \sum_{k=1}^{n} X_k \sim \mathrm{Ga}(n, \lambda).$$

该分布称为参数是n, λ的Γ分布. Γ分布对参数n具有再生性.

- **例(指数分布随机变量的和与差):** 设X 和Y 独立, 均服从指数分布$Exp(\lambda)$, 求1) $Z = X + Y$的概率密度函数.

  2) $Z = X - Y$的概率密度函数(拉普拉斯分布).

(2). 两个随机变量的商

- 如果(X,Y)是二维连续型随机变量，它们的联合密度函数为$f(x,y)$.
- 它们的商$Z=X/Y$是连续型的随机变量，具有密度函数

$$f_{\frac{X}{Y}}(z)=\int_{-\infty}^{+\infty}|t|f(zt,t)dt,\ \forall z\in\mathbb{R}.$$

- 类似的，随机变量Y/X是连续型的随机变量，具有密度函数

$$f_{\frac{Y}{X}}(z)=\int_{-\infty}^{+\infty}|t|f(t,zt)dt,\ \forall z\in\mathbb{R}.$$

- 假若X和Y是独立的随机变量，均服从$N\left(0,\sigma^2\right)$, $Z=X/Y$，则

$$f_Z(z)=\frac{1}{\pi(1+z^2)},$$

称为柯西(Cauchy)分布.

(3). 随机变量的最值

- 若随机变量$X_1, \dots, X_n$相互独立，且对应的分布函数为$F_1(x), \dots, F_n(x)$。
- $X_{(n)} = \max\{X_1, \dots, X_n\}$ 的分布函数？

$$F_{X_{(n)}} = \prod_{i=1}^{n} F_i(x)$$

- $X_{(1)} = \min\{X_1, \dots, X_n\}$的分布函数？

$$F_{X_{(1)}} = 1 - \prod_{i=1}^{n} \left(1 - F_i(x)\right)$$

- 由CDF求PDF/PMF.

例子

- 设 $X_1,\ldots,X_n i.i.d \sim U(0,\theta),\theta>0$, 求 $X_{(n)}=max_{1\leq i\leq n}X_i$ 的密度函数.

- $X_1,\ldots,X_n\sim \mathrm{Exp}(\lambda)$且相互独立，$X_{(1)}=\min\{X_1,\ldots,X_n\}$服从什么分布？

- 掷两颗骰子, 设X和Y分别表示第一和第二次掷出的点数, 求 $P(\max(X,Y)=5)$ 和 $P(\min(X,Y)=3)$.

本章总结

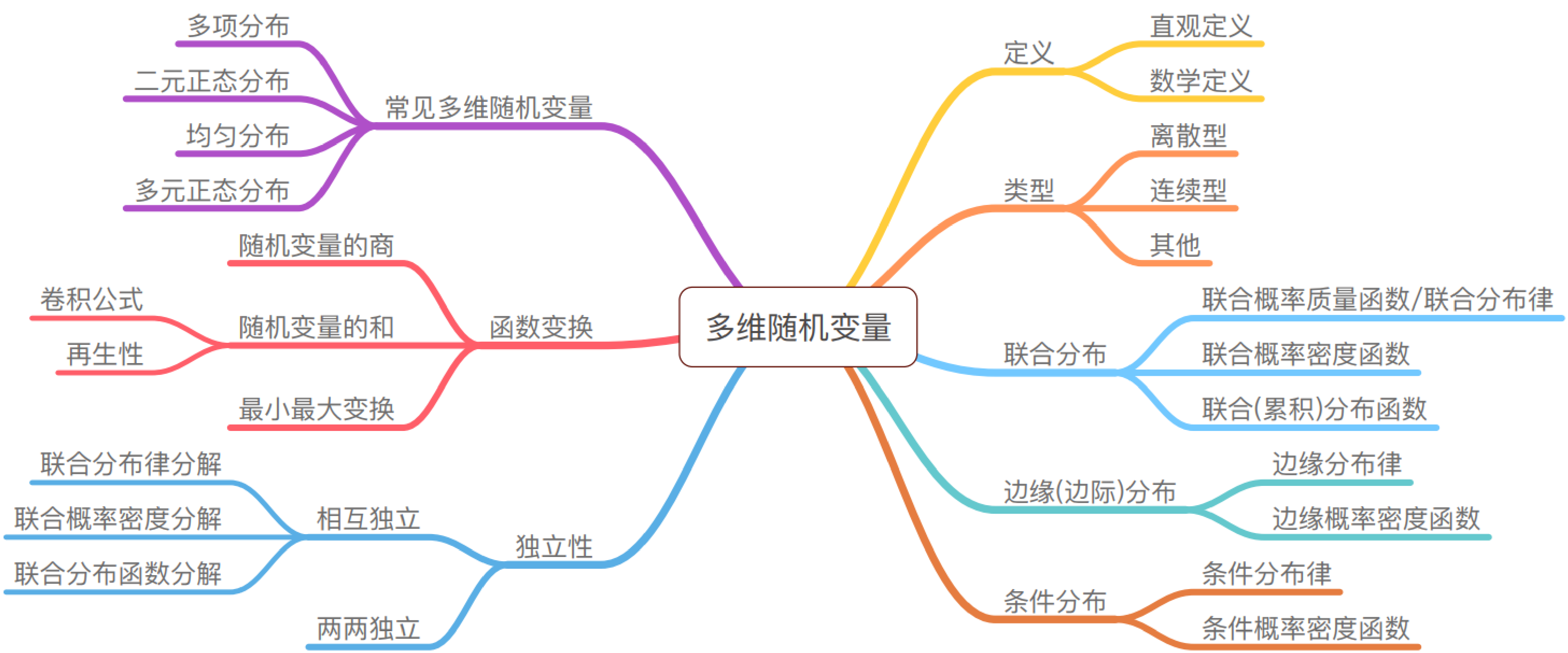

图 3.10: 第三章知识点结构图

Thank you!