二重期望公式

定理 0.30 [二重期望公式] 设 (X,Y) 是二维随机向量, 且 $\mathbb{E}(X)$ 存在, 则

$$\begin{split} &\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) \\ &= \begin{cases} \sum_{i} \mathbb{E}(X|Y=y_{i}) P(Y=y_{i}), & \text{若 } Y \text{ 是一个离散型随机变量} \\ \int_{-\infty}^{+\infty} \mathbb{E}(X|Y=y_{i}) f_{Y}(y) \; dy, & \text{若 } Y \text{ 是一个连续型随机变量} \end{cases} \end{split}$$

二重期望公式在实际中很有用,譬如,在计算取值范围很大的 X 的期望 $\mathbb{E}(X)$ 时,可以通过构建与 X 有关的量 Y,通过 Y 的不同取值将大范围 划分成若干小区域. 先在小区域上求 X 的平均, 再对此类平均求加权平均,即可得到大范围上 X 的期望 $\mathbb{E}(X)$.

推论: 全期望公式

回顾: 对任意事件 A 而言, 根据全概率公式有 $P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$; 即通过对样本空间的切割划分, 将一个复杂事件化成相对简单的事件来求其概率; 借鉴这种化繁为简的思想, 也可以通过全期望公式计算复杂事件的期望, 具体如下:

定理 0.31 设 $A_1, A_2, ..., A_n$ 是样本空间 Ω 一个分割, $A_i A_j = \emptyset$ 和 $\Omega = \bigcup_{i=1}^n A_i$. 对任意随机变量 X 有

$$\mathbb{E}(X) = \mathbb{E}(X|A_1)P(A_1) + \mathbb{E}(X|A_2)P(A_2) + \dots + \mathbb{E}(X|A_n)P(A_n),$$

特别的, 随机事件 A 及其对立事件 \overline{A} 构成空间 Ω 一个分割, 对任意随机 变量 X 有

$$\mathbb{E}(X) = \mathbb{E}(X|A)P(A) + \mathbb{E}(X|\bar{A})P(\bar{A}).$$

可证.

证明: 全期望公式

根据二重期望公式定理 0.30, 假设 Y 有 m 个取值

$$\mathbb{E}(X) = \mathbb{E}_{Y} [\mathbb{E}(X \mid Y)] = \sum_{i=1}^{m} \mathbb{E}(X \mid Y = y_{i}) P(Y = y_{i})$$

$$= \sum_{i=1}^{m} \left[\sum_{j} x_{j} P(x_{j} \mid Y = y_{i}) \right] P(Y = y_{i})$$

$$= \sum_{i=1}^{m} \sum_{j} x_{j} P(x_{j} \mid Y = y_{i}) P(Y = y_{i})$$

$$= \sum_{k=1}^{n} \sum_{j} x_{j} P(x_{j} \mid A_{k}) P(A_{k})$$

最后一步成立于

$$\sum_{k=1}^{n} P(x_j \mid A_k) P(A_k) = \sum_{k=1}^{n} \sum_{k_i : y_{k_i} \in A_k} P(x_j \mid y_{k_i}) P(y_{k_i} \mid A_k) P(A_k)$$
$$= \sum_{i=1}^{m} P(x_j \mid Y = y_i) P(Y = y_i)$$

多维随机向量函数的条件期望:例 0.99

例 0.99 一矿工被困在有三个门的矿井里,第一个门通一坑道,沿此坑道走3小时可使他到达安全地点;第二个门可使他走5小时后回到原处;第三个门可使他走7小时后也回到原地.如设此矿工在任何时刻都等可能地选定其中一个门,试问他到达安全地点平均要用多长时间?

解答:例 0.99

题目: 如上所述.

解答:

 \bullet 设 X 为该矿工回到安全地点所需的时间, 易知 X 的取值为

$$3, 5 + 3, 7 + 3, 5 + 5 + 3, 5 + 7 + 3, 7 + 7 + 3, \dots$$

直接计算取值范围很大的 X 的期望 $\mathbb{E}(X)$ 较困难, 因此构建变量 Y 表示该矿工第一次选择的门的序号. 因此

$$P(Y = 1) = P(Y = 2) = P(Y = 3) = \frac{1}{3},$$

• 若矿工选择第一个门, 则 $\mathbb{E}(X|Y=1)=3$; 若矿工选择第二个门, 则 $\mathbb{E}(X|Y=2)=5+\mathbb{E}(X)$; 若矿工选择第三个门, 则 $\mathbb{E}(X|Y=3)=7+\mathbb{E}(X)$; 综上所述, 由定理0.30有

$$\mathbb{E}(X) = \frac{1}{3}[3 + 5 + \mathbb{E}(X) + 7 + \mathbb{E}(X)] = 5 + \frac{2}{3}\mathbb{E}(X),$$

由此可得 $\mathbb{E}(X) = 15$.

多维随机向量函数的协方差

定义 0.56 设二维随机向量 (X,Y) 的期望 $\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$ 存在,则称其为X 与Y 的协方差,记为

$$Cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

从协方差的定义可以看出, 它是 X 的偏差 " $X - \mathbb{E}[X]$ " 与 Y 的偏差 " $Y - \mathbb{E}[Y]$ " 乘积的数学期望, 由于偏差可正可负, 故协方差也可正可负, 也可为零.

协方差的性质

• 对任意随机变量 X 与 Y, 有

$$Cov(X, X) = 0$$
 $\mathbb{A}\mathbb{R}(X \pm Y) = \mathbb{VAR}(X) + \mathbb{VAR}(Y) \pm 2Cov(X, Y)$

• 对任意随机变量 X 与 Y 和常数 c, 有

$$Cov(X, c) = 0$$
 $TD Cov(X, Y) = Cov(Y, X)$

• 对任意随机变量 X_1 、 X_2 和 Y, 有

$$Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$$

- 若随机变量 X 与 Y 相互独立,则有 Cov(X,Y) = 0,但反之不成立;
- 对任意随机变量 X 与 Y 有

$$(Cov(X,Y))^2 \le VAR(X) \cdot VAR(Y)$$

等号成立的充要条件是 Y = aX + b 几乎处处成立, 即 X 与 Y 之间几乎处处存在 线性关系. (可证)

多维随机向量函数的协方差:例 0.100

例 0.100 设随机变量 (X,Y) 的联合概率密度为

$$f(x,y) = \begin{cases} (x,y)/8, & 0 \le x \le 2, 0 \le y \le 2 \\ 0, & \text{#$\stackrel{}{\mathcal{C}}$} \end{cases}$$

求协方差 Cov(X,Y) 和方差 $\sigma(X+Y)$.

解答:例 0.100

题目: 如上所述.

解答:

• 根据协方差的定义 $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, 计算

$$\mathbb{E}[X] = \mathbb{E}[Y] = \int_0^2 \int_0^2 \frac{x(x+y)}{8} \, dx dy = \frac{7}{6} \,, \quad \mathbb{E}[XY] = \int_0^2 \int_0^2 \frac{xu(x+y)}{8} \, dx dy = \frac{4}{3}$$

由此可得 Cov(X, Y) = -1/36.

• 根据方差的定义 $\sigma(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$, 计算

$$\mathbb{E}[X^2] = \mathbb{E}[Y^2] = \int_0^2 \int_0^2 x^2 (x+y)/8 dx dy = 5/3$$

得 $\sigma(X) = \sigma(Y) = 11/36$. 最后得到

$$\sigma(X+Y) = \sigma(X) + \sigma(Y) + 2\operatorname{Cov}(X,Y) = 5/9.$$

多维随机向量函数的协方差:例 0.101

例 0.101 有 n 对夫妻参加一次聚会,将所有参会人员任意分成 n 组,每组一男一女,用 X 表示夫妻两人被分到一组的对数,求 X 的期望和方差.

解答:例 0.101

题目: 有n 对夫妻参加一次聚会,将所有参会人员任意分成n 组,每组一男一女,用X 表示夫妻两人被分到一组的对数,求X 的期望和方差.

解答:

• 用 X_i 表示第 i 对夫妻是否被分到一组,即

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 对夫妻被分到一组} \\ 0, & \text{否则} \end{cases}$$

则 $X = X_1, X_2, \ldots, X_n$. 随机变量 X_i 得分布列为

$$P(X_i = 1) = (n-1)!/n! = 1/n,$$
 $P(X_i = 0) = 1 - 1/n$

于是得到期望

$$\mathbb{E}[X] = \mathbb{E}(X_1, X_2, \dots, X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) \cdots + \mathbb{E}(X_n) = 1.$$

• 对任意 $i \neq j$,有

$$P(X_i = 1, X_j = 1) = (n-2)!/n! = 1/n(n-1),$$

由此得到

$$Cov(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \mathbb{E}[X_j] = 1/n^2(n-1),$$

最后根据协方差的性质有

$$\sigma(X) = \sum_{i=1}^{n} \sigma(X_i) + 2\sum_{i \neq j} \operatorname{Cov}(X_i, X_j) = 1.$$

插播:线性运算的基本性质

•回忆: 矩阵的特征值和特征向量. 以二维为例, 矩阵 \mathbf{A} 拥有特征值 λ_1 和 λ_2 , 分别对应特征向量 $\mathbf{u}_1, \mathbf{u}_2$. 则有

$$\mathbf{A}\mathbf{u}_i = \mathbf{\Lambda}_i \mathbf{u}_i \;, \quad i = 1, 2$$

进而有 $\mathbf{A} = \mathbf{U}^{\mathsf{T}} \mathbf{\Lambda} \mathbf{U}$ with $\mathbf{U}^{\mathsf{T}} = \mathbf{U}^{-1}$, 即

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} \begin{pmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$$

- ullet 几何性质. For the case of ${f A}{m v}$ given any ${m v}$,
 - •矩阵 U 负责对 v 进行旋转
 - •矩阵 Λ 负责对 v 进行放缩

插播:线性运算的基本性质

如果我们面对一个运算 $\mathbf{A}x + \mathbf{b}$,

- $\bullet b$ 是平移
- U 是旋转
- Λ 是放缩

二维正态分布的协方差

定理 0.32 若随机向量 $(X,Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, 则

$$Cov(X,Y) = \rho \, \sigma_x \sigma_y$$

推论 若随机向量 $(X,Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, 则 X 与 Y 相互独立的 充要条件是 Cov(X,Y) = 0.

可证.

证明:二维正态分布的协方差

二维正态分布的密度函数为

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_x\sigma_y} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x^2\sigma_y^2} \right] \right)$$

这里用到了坐标变换(归一化做法)

$$f(u,v) = f(x,y) |\mathbf{J}|, \quad (u,v) = \left(\frac{x - \mu_x}{\sigma_x}, \frac{y - \mu_y}{\sigma_y}\right), \quad |\mathbf{J}| = \sigma_x \sigma_y.$$

则有

$$Cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_x)(y - \mu_y) f(x,y) \, dx dy$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\sigma_x \sigma_y}{2\pi \sqrt{1 - \rho^2}} \, uv \exp\left(-\frac{u^2 + v^2 - 2\rho uv}{2(1 - \rho^2)}\right) \, du dv$$

$$= \frac{\sigma_x \sigma_y}{2\pi \sqrt{1 - \rho^2}} \int_{-\infty}^{+\infty} v \exp\left(-\frac{v^2}{2}\right) \left[\int_{-\infty}^{+\infty} u \exp\left(-\frac{(u - \rho v)^2}{2(1 - \rho^2)}\right) du\right] dv$$

$$= \frac{\sigma_x \sigma_y}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \rho v^2 \exp\left(-\frac{v^2}{2}\right) dv = \rho \sigma_x \sigma_y$$

协方差与方差

• 方差. 衡量单变量自身的波动性或者偏离性.

$$\mathbb{VAR}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

• 协方差. 衡量变量间的偏离性

$$Cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

可以定义矩阵∑用以衡量多变量的偏离程度

$$\Sigma = \begin{pmatrix} \mathbb{VAR}(X) & \operatorname{Cov}(X,Y) \\ \operatorname{Cov}(X,Y) & \mathbb{VAR}(Y) \end{pmatrix}$$

随机向量的数学期望与协方差阵

n 维随机向量的数学期望及方差可以通过矩阵形式给出.

定义 **0.57** 设 n 维随机向量为 $X = (X_1, X_2, ..., X_n)'$, 若每个分量的数学期望都存在, 则称

$$\mathbb{E}[\mathbf{X}] = (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_n])',$$

为X的数学期望向量,简称X的数学期望.而称

$$\mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])']$$

$$= \begin{bmatrix} \sigma_{X_1}^2 & \operatorname{Cov}(X_1, X_2) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & \sigma_{X_2}^2 & \cdots & \operatorname{Cov}(X_2, X_n) \\ \vdots & \vdots & & \vdots \\ \operatorname{Cov}(X_n, X_1) & \operatorname{Cov}(X_n, X_2) & \cdots & \sigma_{X_n}^2 \end{bmatrix}$$

为X的方差-协方差矩阵,简称X的协方差阵.

随机向量的协方差阵的性质

通过定义 0.57 可以看到 n 维随机向量的各分量的方差构成了协方差阵对角线上的元素, 非对角线的元素为协方差.

定理 0.33 n 维随机向量的协方差阵 $Cov(\mathbf{X}) = (Cov(X_i, X_j))_{n \times n}$ 是一个对称的非负定矩阵.

Remarks:

- •这说明, 协方差矩阵的特征值是实数的、非负的.
- •多维随机向量函数的协方差:例 0.102

例 0.102 设随机变量 $X_1, X_2, ..., X_n$ 相互独立且服从正态分布, 方差为 σ^2 . 记 $\bar{X} = \sum_{i=1}^n X_i/n$, 讨论 \bar{X} 和 $\bar{X} - X_i$ 的独立性.

解答: 例 0.102

题目:设随机变量 X_1, X_2, \ldots, X_n 相互独立且服从正态分布, 方差为 σ^2 . 记 $\bar{X} = \sum_{i=1}^n X_i/n$, 讨论 \bar{X} 和 $\bar{X} - X_i$ 的独立性.

解答:

• 根据正态分布的性质易知 \bar{X} 和 $\bar{X} = X_i$ 都服从正态分布 (线性性), 根据定理 0.32 可知正态分布的独立性可通过协方差来研究. 根据协方差的性质有

$$\operatorname{Cov}(\bar{X}, \bar{X} - X_i) = \operatorname{Cov}(\bar{X}, \bar{X}) - \operatorname{Cov}(\bar{X}, X_i) = \sigma(\bar{X}) - \operatorname{Cov}\left(\sum_{j=1}^n \frac{X_j}{n}, X_i\right)$$

• 根据 X_1, X_2, \ldots, X_n 相互独立有

$$\sigma(\bar{X}) = \frac{1}{n^2} \sigma\left(\sum_{i=1}^n X_i\right) = \frac{\sigma^2}{n} \qquad \text{fil} \qquad \operatorname{Cov}(\sum_{j=1}^n \frac{X_j}{n}, X_i) = \frac{1}{n} \operatorname{Cov}(X_i, X_i) = \frac{\sigma^2}{n}$$

于是得到 $Cov(\bar{X}, \bar{X} - X_i) = 0$, 根据定理 0.32 的推论可知 \bar{X} 和 $\bar{X} - X_i$ 相互独立.

多维随机向量函数的相关系数

两个随机变量之间的关系可以分为独立与非独立,其中非独立关系中又可以分为线性关系和非线性关系,线性相关程度通过线性相关系数来定义.

定义 0.58 设 (X,Y) 为二维随机向量, 如果它们的标准差 σ_x 和 σ_y 存在且都不为零, 则称

$$\rho_{XY} = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\mathbb{VAR}(X) \, \mathbb{VAR}(Y)}} .$$

为X与Y的线性相关系数,简称相关系数.

相关系数的性质

- •对任意随机变量 X 与 Y, 有 $|\rho_{XY}| \le 1$. 等号成立的充要条件是 Y = aX + b 几乎处处成立, 即 X 与 Y 之间几乎处处存在线性关系.
 - •若 $\rho_{XY} = 0$, 称 X 与 Y 不相关. 不相关是指 X 与 Y 之间没有线性关系, 但 X 与 Y 之间可能存在其他的函数关系, 比如平方关系、对数关系等;
 - •若 $\rho_{XY} = 1$, 称 X 与 Y 完全正相关; 若 $\rho_{XY} = -1$, 称 X 与 Y 完全 负相关;
 - •若 $0 < |\rho_{XY}| < 1$,称 X 与 Y "有一定程度" 的线性关系; 若 $|\rho_{XY}|$ 越接近于 1, 则线性相关程度越高; 若 $|\rho_{XY}|$ 越接近于 0, 则线性相关程度越低.
- •若随机变量 X 与 Y 相互独立,则 X 与 Y 不相关,但反之不成立.

正态分布的相关系数

定理 0.34 若随机向量 $(X,Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, 则有

- $X \to Y$ 的线性相关系数 $\rho_{XY} = \rho$
- X 与 Y 相互独立充要条件是 X 与 Y 不相关, 即 $\rho = 0$.

Remarks: 独立与不相关的等价性仅限于正态分布随机变量, 对于其他类型不一定成立.

不相关的等价条件

定理 0.35 若随机变量 X 与 Y 的方差存在且都不为零,以下几个条件相互等价:

- $\bullet X$ 与Y独立;
- X 与 Y 不相关, 即 $\rho_{XY} = 0$;
- 协方差 Cov(X,Y) = 0;
- $\bullet \mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y] ;$
- $\bullet \, \mathbb{VAR}(X \pm Y) = \mathbb{VAR}(X) + \mathbb{VAR}(Y) \ .$

多维随机向量函数的相关系数:例 0.103

例 0.103 设随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$ 和 $Y \sim \mathcal{N}(\mu, \sigma^2)$ 相互独立, 求 $Z_1 = \alpha X + \beta Y$ 和 $Z_2 = \alpha X - \beta Y$ 的相关系数 $(\alpha, \beta \neq 1)$.

解答:例 0.103

题目: 设随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$ 和 $Y \sim \mathcal{N}(\mu, \sigma^2)$ 相互独立, 求 $Z_1 = \alpha X + \beta Y$ 和 $Z_2 = \alpha X - \beta Y$ 的相关系数 $(\alpha, \beta \neq 1)$.

解答:

• 根据相关系数的定义

$$\rho_{Z_1 Z_2} = \frac{\operatorname{Cov}(Z_1, Z_2)}{\sigma_{Z_1} \sigma_{Z_2}}$$

计算

$$Cov(Z_1, Z_2) = Cov(\alpha X + \beta Y, \alpha X - \beta Y) = (\alpha^2 - \beta^2)\sigma^2$$

$$\sigma_{Z_1}^2 = Cov(\alpha X + \beta Y, \alpha X + \beta Y) = (\alpha^2 + \beta^2)\sigma^2$$

$$\sigma_{Z_2}^2 = Cov(\alpha X - \beta Y, \alpha X - \beta Y) = (\alpha^2 + \beta^2)\sigma^2$$

由此可得

$$\rho_{Z_1 Z_2} = \frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2} \ .$$

多维随机向量函数的相关系数:例 0.104

例 0.104 设随机向量 $(X_1, X_2, ..., X_n)$ 服从多项分布 $M(m, p_1, p_2, ..., p_n)$, 对任意 $i \neq j$, 求 X_i 和 X_j 的相关系数.

解答: 例 0.104

题目: 设随机向量 $(X_1, X_2, ..., X_n)$ 服从多项分布 $M(m, p_1, p_2, ..., p_n)$, 对任意 $i \neq j$, 求 X_i 和 X_j 的相关系数.

解答:

•根据多项分布的性质,有边缘分布

$$X_i \sim B(m, p_i)$$
 π $X_j \sim B(m, p_j)$

由此可得 $\sigma(X_i) = mp_i(1 - p_i)$ 和 $\sigma(X_j) = mp_j(1 - p_j)$.

• 对每个 $k \in [m]$, 引入随机变量

$$Y_i^k = \begin{cases} 1, & \text{若第 } k \text{ 次实验的结果为 } i \\ 0, & \text{其它} \end{cases} \qquad \text{和} \quad Y_j^k = \begin{cases} 1, & \text{若第 } k \text{ 次实验的结果为 } j \\ 0, & \text{其它} \end{cases}$$

由此可得

$$X_i = Y_i^1 + Y_i^2 + \dots + Y_i^m$$
 $X_j = Y_j^1 + Y_j^2 + \dots + Y_j^m$

J20

•根据第 k 次实验和第 l 次实验相互独立 $(k \neq l)$, 以及 $Y_i^k Y_i^{k=0}$ 有

$$\operatorname{Cov}(Y_i^k, Y_j^k) = 0 \qquad \text{fil} \qquad \operatorname{Cov}(Y_i^k, Y_j^k) = \mathbb{E}[Y_i^k Y_j^k] - \mathbb{E}[Y_i^k] \mathbb{E}[Y_j^k] = -p_i p_j$$

根据协方差的性质有

$$Cov(X_i, X_j) = \sum_{k=1}^m Cov(Y_i^k, Y_j^k) + \sum_{k \neq l} Cov(Y_i^k, Y_j^l) = -mp_i p_j$$

由此可得 X_i 和 X_i 的相关系数

$$\rho = \frac{\text{Cov}(X_i, X_j)}{\sigma(X_i)\sigma(X_j)} = \frac{-mp_i p_j}{\sqrt{mp_i(1 - p_i)}\sqrt{mp_j(1 - p_j)}} = -\frac{p_i p_j}{\sqrt{p_i(1 - p_i)}\sqrt{p_j(1 - p_j)}}$$

二维正态分布的相关总结

二维随机变量 $(X,Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, 则随机变量 X 和 Y 的边缘分布为

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_x\sigma_y} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x^2\sigma_y^2} \right] \right)$$

- 边缘分布服从正态分布 $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ 和 $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$
- 条件分布服从正态分布

$$(X \mid Y = y) \sim \mathcal{N}\left(\mu_x + \rho(y - \mu_y)\frac{\sigma_x}{\sigma_y}, (1 - \rho^2)\sigma_x^2\right)$$

• 正态分布之和是正态分布

$$X + Y \sim \mathcal{N}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

- 协方差 $Cov(X,Y) = \rho \sigma_x \sigma_y$
- 相关系数

$$\rho = \rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}$$

• X 与 Y 独立, 充要条件, Cov(X,Y) = 0 或者 $\rho = \rho_{XY} = 0$