

概 率 论

Probability

已学知识点

● 第一章 事件与概率

▶ 随机现象与统计规律性

- ① 概率的频率解释依然是当今最通行的解释.
- ② 描述频率趋近于概率的大数定律总是概率论的第一大数定律.
- ③ 实际当中用频率作为概率的估计是十分自然的.

▶ 样本空间与事件

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω
$A \subset B$	A 是 B 的子集	A 发生则 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A - B$	差集	A 发生而 B 不发生
\bar{A}	余集	A 不发生

已学知识点

● 第一章 事件与概率

- ▶ 古典概型 (等可能概率模型): (1) 样本空间样本点有限; (2) 每个样本点等可能出现.
 - 计数方法: 排列组合.
 - 三个基本性质: 非负性、规范性、有限可加性.
- ▶ 几何概率: 以等可能性定义概率, 处理无限场合, 概率是几何体的测度之比.
 - 三个基本性质: 非负性、规范性、可列可加性.
- ▶ 概率空间: (Ω, \mathcal{F}, P)
 - 难点和要点: 事件域 \mathcal{F} 的选择, 太小不能满足需要, 太大难以定义概率.
 - 选择包含我们关注的所有事件的 σ 域, 保证事件对交、并、逆、差作可列次运算的封闭性.
 - 在这种 σ 域上, 能定义满足非负、规范和可列可加性的概率测度.

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 条件概率: $P(A|B) = \frac{P(AB)}{P(B)}$.

- 乘法公式: $P(A_1A_2\cdots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1A_2) \cdot \cdots \cdot P(A_n|A_1A_2\cdots A_{n-1})$

- 全概率公式: $P(B) = \sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i)$

- Bayes 公式: $P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i)}$

$$\left. \begin{array}{l} A_i \cap A_j = \Phi \quad (i \neq j) \\ \bigcup_{i=1}^{\infty} A_i = \Omega \end{array} \right\}$$

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 事件独立性：两个事件 $P(AB) = P(A) \cdot P(B)$. 三个事件
$$\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$$

- ▶ 试验独立性：一个试验的结果对其它各试验的可能结果的概率都无影响.

- ▶ Bernoulli 试验 E : 概率空间 (Ω, \mathcal{F}, P) , 其中

$$A \subset \Omega, \quad \mathcal{F} = \{\Phi, A, \bar{A}, \Omega\}, \quad P(A) = p, \quad P(\bar{A}) = q, \quad (p > 0, q > 0, p + q = 1)$$

- Bernoulli 分布

- 二项分布

- 几何分布

- Pascal 分布

- 多项分布

- Poisson 分布

已学知识点

● 第三章 随机变量与分布函数

▶ 随机变量 (r.v.) 与分布函数 (c.d.f.):

- 随机变量 ξ 是概率空间 (Ω, \mathcal{F}, P) 中定义域为 Ω 、取值于 \mathbb{R} 的一个单值 Borel 函数.
- 分布函数 $F(x) = P\{\xi(\omega) < x\}$, $-\infty < x < \infty$ 是单调不降、取值 $[0, 1]$ 的左连续函数. 它完整描述了随机变量, 是研究的主要对象.
- 随机变量依取值 $\begin{cases} \text{离散型:} & \text{分布律 (mass function)、分布列} \\ \text{连续型:} & \text{概率密度 (p.d.f.)} \end{cases}$
- 主要分布: $\begin{cases} \text{离散型:} & \text{Bernoulli, binomial, Poisson, hyper-geometric, geometric} \\ \text{连续型:} & \text{uniform, exponential, normal, } \Gamma \end{cases}$
- 概率计算: $\begin{cases} \text{离散型:} & P\{x \in D\} = \sum_{x_i \in D} p_i, \quad P\{(x, y) \in D\} = \sum_{(x_i, y_j) \in D} p_{ij} \\ \text{连续型:} & P\{x \in D\} = \int_D p(x) dx, \quad P\{(x, y) \in D\} = \iint_D p(x, y) dx dy \end{cases}$

已学知识点

● 第三章 随机变量与分布函数

▶ 随机向量，随机变量的独立性：

○ 随机向量即多元随机变量

{	联合分布： 联合分布函数、联合分布律、联合密度 边际分布： 边际分布函数、边际分布律、边际密度 条件分布： 条件分布函数、条件分布律、条件密度 独立性： 与事件独立性几乎完全相同
---	--

○ 主要分布：

{	离散型： 多项分布、多元超几何分布 连续型： 二元均匀分布、二元正态分布
---	---

$$(\xi, \eta) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}\right) \implies \left\{ \begin{array}{l} \xi \sim N(\mu_1, \sigma_1^2), \quad \eta \sim N(\mu_2, \sigma_2^2) \\ (\eta | \xi = x) \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2)\right) \\ (\xi | \eta = y) \sim N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2), \sigma_1^2 (1 - \rho^2)\right) \\ \xi \text{ 与 } \eta \text{ 相互独立} \iff \rho = 0 \end{array} \right.$$

已学知识点

● 第三章 随机变量与分布函数

▶ 随机变量的函数及其分布:

○ 随机变量的函数什么情况下还是随机变量

○ 离散型易: $\begin{cases} \eta = g(\xi) : & \text{对应法} \\ \zeta = g(\xi, \eta) : & \text{表上作业法, 独立情形和的卷积公式} \end{cases}$

○ 连续型难: $\begin{cases} \eta = g(\xi) : \text{直接法 } F_{\eta}(y) = P\{\eta < y\} = \int_{g(x) < y} p_{\xi}(x) dx \\ \zeta = g(\xi, \eta) : \text{直接法 } F_{\zeta}(z) = P\{\zeta < z\} = \iint_{g(x, y) < z} p(x, y) dx dy \\ \text{和 (卷积公式)、差、商、积、max, min 的公式} \end{cases}$

$\begin{cases} \zeta_1 = g_1(\xi, \eta) \\ \zeta_2 = g_2(\xi, \eta) \end{cases} : \text{变换法 } \begin{cases} u = g_1(x, y) \\ v = g_2(x, y) \end{cases} \implies \begin{cases} x = h_1^{-1}(u, v) \\ y = h_2^{-1}(u, v) \end{cases} \implies J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$

$$q_{\zeta_1, \zeta_2}(u, v) = p(h_1^{-1}(u, v), h_2^{-1}(u, v)) \cdot |J|$$

已学知识点

● 第四章 数字特征与特征函数

▶ 数学期望

○ 定义:
$$\left\{ \begin{array}{l} \text{离散型: } E(\xi) = \sum_i x_i \cdot p_i \\ \text{连续型: } E(\xi) = \int_{-\infty}^{\infty} x \cdot p_{\xi}(x) dx \end{array} \right.$$

绝对收敛!

○ 随机变量函数的数学期望:
$$\left\{ \begin{array}{l} \text{离散型: } E[g(\xi)] = \sum_i g(x_i) \cdot p_i \\ \text{连续型: } E[g(\xi)] = \int_{-\infty}^{\infty} g(x) \cdot p_{\xi}(x) dx \end{array} \right.$$

$$\left\{ \begin{array}{l} \text{离散型: } E[g(\xi, \eta)] = \sum_i \sum_j g(x_i, y_j) \cdot p_{ij} \\ \text{连续型: } E[g(\xi, \eta)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot p_{(\xi, \eta)}(x, y) dx dy \end{array} \right.$$

已学知识点

- 第四章 数字特征与特征函数

- ▶ 数学期望

- 数学期望的性质：
 - ① $E(c) = c$
 - ② $a \leq \xi \leq b \implies a \leq E(\xi) \leq b$
 - ③ $\xi \leq \eta \implies E(\xi) \leq E(\eta)$
 - ④ $E(a\xi + b\eta) = a \cdot E(\xi) + b \cdot E(\eta)$
 - ⑤ ξ, η 独立 $\implies E(\xi\eta) = E(\xi) \cdot E(\eta)$

已学知识点

● 第四章 数字特征与特征函数

▶ 方差: $D(\xi) = E\left\{\left[\xi - E(\xi)\right]^2\right\}$ 标准差: $\sqrt{D(\xi)}$ 计算公式: $D(\xi) = E(\xi^2) - [E(\xi)]^2$

○ 方差的性质: ① 设 C 是常数, 则 $D(C) = 0$.

② 设 ξ 是随机变量, c 为常数, 则 $D(\xi + c) = D(\xi)$.

③ 设 ξ 是随机变量, c 为常数, 则 $D(c\xi) = c^2D(\xi)$.

④ 若 $c \neq E(\xi)$, 则 $D(\xi) < E[(\xi - c)^2]$.

⑤ $D(\xi) = 0 \iff P\{\xi = C\} = 1$.

⑥ 若 ξ 与 η 独立, 则 $D(\xi \pm \eta) = D(\xi) + D(\eta)$.

已学知识点

● 第四章 数字特征与特征函数

▶ 常见分布的数学期望与方差:

分布	符号表示	数学期望	方差
二项分布	$\xi \sim b(n, p)$	$E(\xi) = np$	$D(\xi) = np(1-p)$
Bernoulli 分布	$\xi \sim b(1, p)$	$E(\xi) = p$	$D(\xi) = p(1-p)$
Poisson 分布	$\xi \sim P(\lambda)$	$E(\xi) = \lambda$	$D(\xi) = \lambda$
均匀分布	$\xi \sim U[a, b]$	$E(\xi) = \frac{a+b}{2}$	$D(\xi) = \frac{(b-a)^2}{12}$
Gamma 分布	$X \sim \Gamma(r, \lambda)$	$E(\xi) = \frac{r}{\lambda}$	$D(\xi) = \frac{r}{\lambda^2}$
指数分布	$\xi \sim \text{Exp}(\lambda)$	$E(\xi) = \frac{1}{\lambda}$	$D(\xi) = \frac{1}{\lambda^2}$
χ^2 分布	$X \sim \chi_n^2$	$E(\xi) = n$	$D(\xi) = 2n$
正态分布	$\xi \sim N(\mu, \sigma^2)$	$E(\xi) = \mu$	$D(\xi) = \sigma^2$

已学知识点

- 第四章 数字特征与特征函数

- ▶ **Chebyshev 不等式:** $P \left\{ \left| \xi - E(\xi) \right| \geq \varepsilon \right\} \leq \frac{D(\xi)}{\varepsilon^2} \iff P \left\{ \left| \xi - E(\xi) \right| < \varepsilon \right\} \geq 1 - \frac{D(\xi)}{\varepsilon^2}$

$$3\sigma \text{ 法则: } P \left\{ \left| \xi - \mu \right| < \varepsilon \right\} \geq 1 - \frac{\sigma^2}{\varepsilon^2} \implies \begin{cases} P \left\{ \left| \xi - \mu \right| \leq \sigma \right\} \geq 0, & \varepsilon = \sigma \\ P \left\{ \left| \xi - \mu \right| \leq 2\sigma \right\} \geq 0.75, & \varepsilon = 2\sigma \\ P \left\{ \left| \xi - \mu \right| \leq 3\sigma \right\} \geq 0.8889, & \varepsilon = 3\sigma \end{cases}$$

4.2 方差，相关系数，矩

一、方差

二、切比雪夫 (Chebyshev) 不等式

三、相关系数

四、矩

五、条件数学期望

三、相关系数

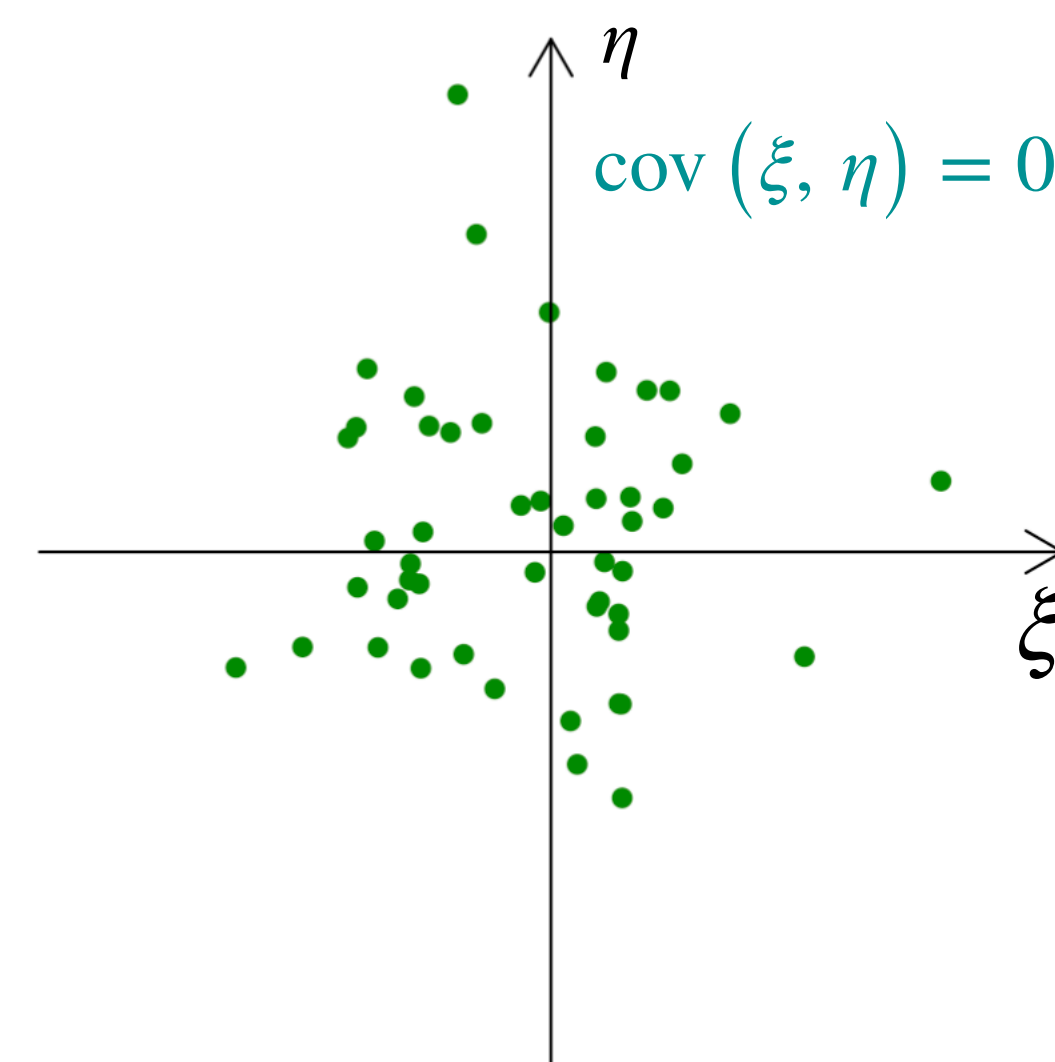
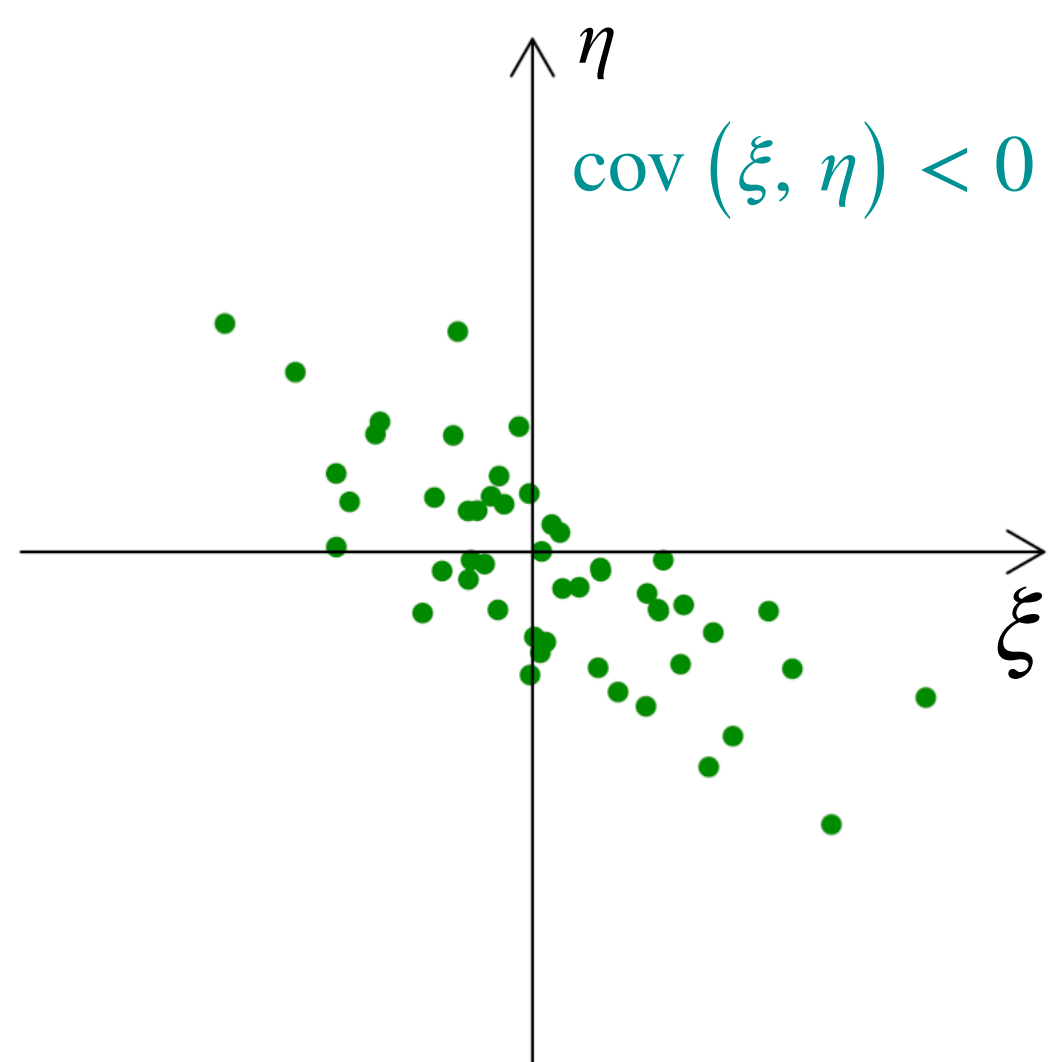
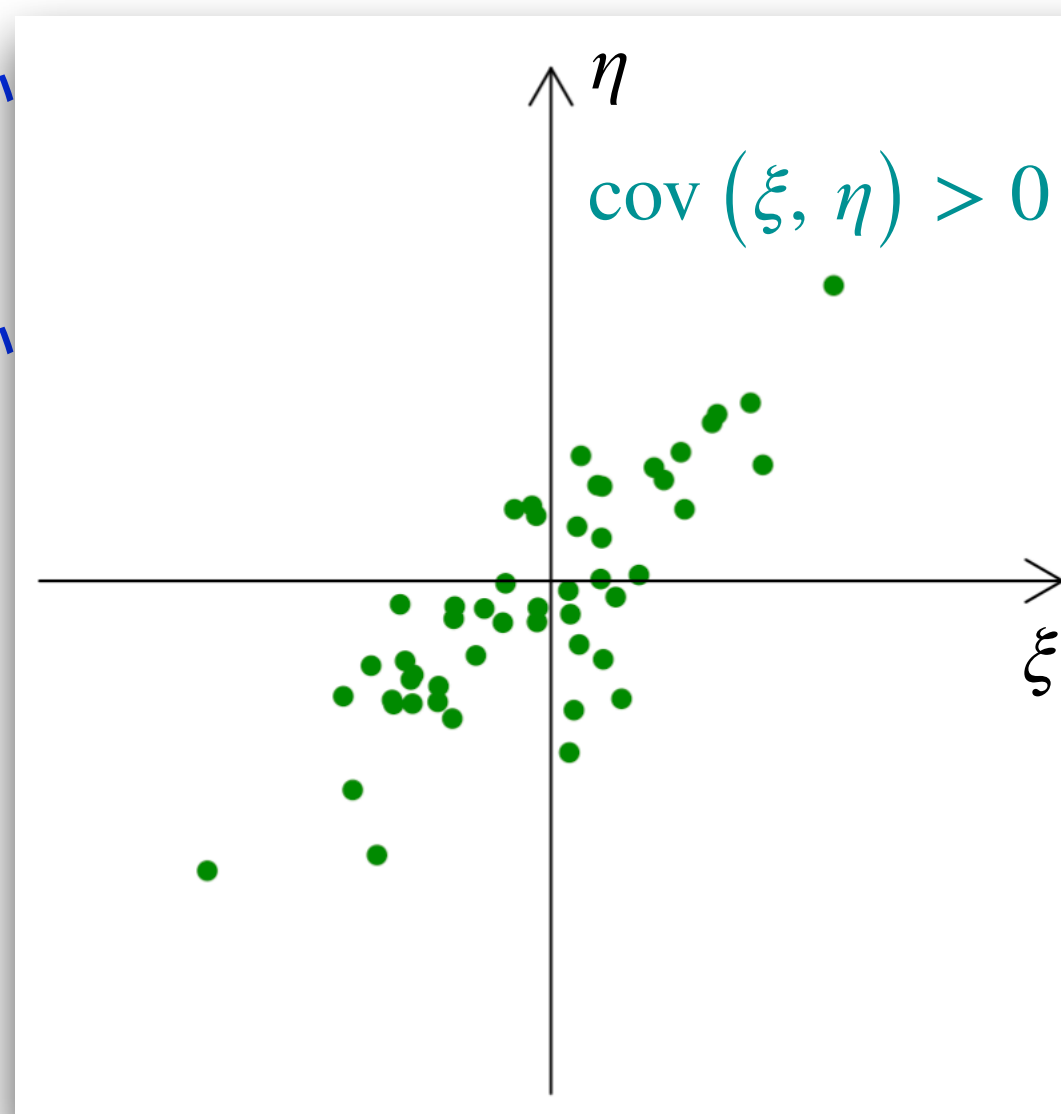
- 对于随机向量, 我们除了关心它的各个分量的概率分布外, 还希望知道各个分量之间的联系.

定义: 设 (ξ, η) 是二维随机向量, 称

$$\text{cov}(\xi, \eta) \triangleq E \left\{ \left[\xi - E(\xi) \right] \left[\eta - E(\eta) \right] \right\}$$

为随机变量 ξ 与 η 的协方差 (covariance).

- 协
- 协



三、相关系数

$$\text{cov}(\xi, \eta) \triangleq E \left\{ \left[\xi - E(\xi) \right] \left[\eta - E(\eta) \right] \right\}$$

- 协方差的性质

① $\text{cov}(X, X) = D(X).$

② $\text{cov}(X, Y) = \text{cov}(Y, X).$

③ $\text{cov}(aX, bY) = ab \cdot \text{cov}(X, Y)$, 其中 a, b 为常数.

④ $\text{cov}(X_1 + X_2, Y) = \text{cov}(X_1, Y) + \text{cov}(X_2, Y).$

⑤ 若 X, Y 相互独立, 则 $\text{cov}(X, Y) = 0.$

⑥ 若 C 为常数, 则 $\text{cov}(X, C) = 0.$

三、相关系数

- 和的方差公式

① $D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y).$

$$\begin{aligned} D(X - Y) &= E \left\{ [(X - Y) - E(X - Y)]^2 \right\} = E \left\{ [(X - E(X)) - (Y - E(Y))]^2 \right\} \\ &= E \left[(X - E(X))^2 \right] + E \left[(Y - E(Y))^2 \right] - 2E \left[(X - E(X))(Y - E(Y)) \right] \\ &= D(X) + D(Y) - 2\text{cov}(X, Y) \end{aligned}$$

② 推广:

$$\begin{aligned} D \left(\sum_{i=1}^n X_i \right) &= \sum_{i,j} \text{cov}(X_i, X_j) = \sum_{i=1}^n D(X_i) + \sum_{i \neq j} \text{cov}(X_i, X_j) \\ &= \sum_{i=1}^n D(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j) \end{aligned}$$

三、相关系数

- 协方差矩阵: 对 n 维随机向量 $\xi = (\xi_1, \xi_2, \dots, \xi_n)$, 记

$$\sigma_{ij} = \text{COV}(\xi_i, \xi_j), \quad i, j = 1, 2, \dots, n$$

则称矩阵

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

为随机向量 ξ 的协方差矩阵 (covariance matrix), 记为 $D(\xi)$.

- ▶ 显然 Σ 是一个对称矩阵.
- ▶ 协方差矩阵 Σ 是一个非负定矩阵. 可以证明, $\forall \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ 都有

$$\mathbf{x} \Sigma \mathbf{x}^T = \sum_{i,j} \sigma_{ij} x_i x_j = \sum_{i,j} \text{COV}(x_i \xi_i, x_j \xi_j) = D\left(\sum_{i=1}^n x_i \xi_i\right) \geq 0 \quad \implies \quad \det \Sigma = |\Sigma| \geq 0$$

三、相关系数

- 协方差的值受量纲影响.

```
library(alr4)
?Heights
var(Heights$mheight, Heights$dheight) # 母亲身高 (英尺) 与女儿身高 (英尺) 的协方差
x = 0.3048 * Heights$mheight # 英尺换算为米
y = 0.3048 * Heights$dheight # 英尺换算为米
var(x, y) # 母亲身高 (米) 与女儿身高 (米) 的协方差
```

Pearson-Lee data

Description

Karl Pearson organized the collection of data on over 1100 families in England in the period 1893 to 1898. This particular data set gives the Heights in inches of mothers and their daughters, with up to two daughters per mother. All daughters are at least age 18, and all mothers are younger than 65. Data were given in the source as a frequency table to the nearest inch. Rounding error has been added to remove discreteness from graph.

Format

This data frame contains the following columns:

mheight

Mother's ht, in.

dheight

Daughter's ht, in.

```
> var(Heights$mheight, Heights$dheight) # 母亲身高 (英尺) 与女儿身高 (英尺) 的协方差
[1] 3.004806
>
> var(x, y) # 母亲身高 (米) 与女儿身高 (米) 的协方差
[1] 0.2791556
```

三、相关系数

定义: 称

$$\rho \triangleq \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$$

为随机变量 ξ 与 η 的相关系数 (correlation coefficient), 其中 $D(\xi) > 0$, $D(\eta) > 0$.

► ξ 与 η 的相关系数是其标准化的随机变量 $\xi^* = \frac{\xi - E(\xi)}{\sqrt{D(\xi)}}$ 与 $\eta^* = \frac{\eta - E(\eta)}{\sqrt{D(\eta)}}$ 的协方差.

$$\text{cov}(\xi^*, \eta^*) = \text{cov}\left(\frac{\xi - E(\xi)}{\sqrt{D(\xi)}}, \frac{\eta - E(\eta)}{\sqrt{D(\eta)}}\right) = \frac{\text{cov}(\xi - E(\xi), \eta - E(\eta))}{\sqrt{D(\xi)}\sqrt{D(\eta)}} = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}} = \rho$$

三、相关系数

- 例: 设 $(\xi_1, \xi_2, \dots, \xi_r)$ 服从多项分布, 求各分量之间的相关系数.

$$P\{\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_r = k_r\} = \frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}, \quad \text{整数 } k_i \geq 0 \text{ 且 } k_1 + k_2 + \dots + k_r = n$$

$$\implies \xi_i \sim b(n, p_i), \quad i = 1, 2, \dots, r \implies E(\xi_i) = np_i, \quad D(\xi_i) = np_i(1 - p_i), \quad i = 1, 2, \dots, r$$

由多项分布的构成 $\implies \xi_i + \xi_j \sim b(n, p_i + p_j), \quad i \neq j \implies \begin{cases} E(\xi_i + \xi_j) = n(p_i + p_j) \\ D(\xi_i + \xi_j) = n(p_i + p_j)(1 - p_i - p_j) \end{cases}$

$$D(\xi_i + \xi_j) = D(\xi_i) + D(\xi_j) + 2\text{cov}(\xi_i, \xi_j) = np_i(1 - p_i) + np_j(1 - p_j) + 2\text{cov}(\xi_i, \xi_j)$$

$$\implies \text{cov}(\xi_i, \xi_j) = -np_i p_j \implies \rho_{ij} = \frac{\text{cov}(\xi_i, \xi_j)}{\sqrt{D(\xi_i)}\sqrt{D(\xi_j)}} = \frac{-np_i p_j}{\sqrt{np_i(1 - p_i)}\sqrt{np_j(1 - p_j)}} = -\sqrt{\frac{p_i p_j}{(1 - p_i)(1 - p_j)}}$$

三、相关系数

- 相关系数的值不受量纲的影响.

```
library(alr4)
?Heights
cor(Heights$mheight, Heights$dheight) # 母亲身高 (英尺) 与女儿身高 (英尺) 的相关系数
x = 0.3048 * Heights$mheight # 英尺换算为米
y = 0.3048 * Heights$dheight # 英尺换算为米
cor(x, y) # 母亲身高 (米) 与女儿身高 (米) 的相关系数
```

Pearson-Lee data

Description

Karl Pearson organized the collection of data on over 1100 families in England in the period 1893 to 1898. This particular data set gives the Heights in inches of mothers and their daughters, with up to two daughters per mother. All daughters are at least age 18, and all mothers are younger than 65. Data were given in the source as a frequency table to the nearest inch. Rounding error has been added to remove discreteness from graph.

Format

This data frame contains the following columns:

mheight

Mother's ht, in.

dheight

Daughter's ht, in.

```
> cor(Heights$mheight, Heights$dheight) # 母亲身高 (英尺) 与女儿身高 (英尺) 的相关系数
[1] 0.4907094
>
> cor(x, y) # 母亲身高 (米) 与女儿身高 (米) 的相关系数
[1] 0.4907094
```

三、相关系数

- 相关系数的性质.

定理: (Cauchy-Schwarz 不等式)

对任意随机变量 ξ 与 η , 如果 $E(\xi^2) < +\infty$, $E(\eta^2) < +\infty$, 则有

$$\left| E(\xi\eta) \right|^2 \leq E(\xi^2) \cdot E(\eta^2)$$

其中等式成立当且仅当存在某一常数 c 使得 $P\{\eta = c\xi\} = 1$ 成立.

$$\xrightarrow{\forall t \in \mathbb{R}} f(t) = E\left[(t\xi - \eta)^2\right] = \left[E(\xi^2)\right] t^2 - 2E(\xi\eta) t + E(\eta^2) \geq 0$$

$$\implies \left[2E(\xi\eta)\right]^2 - 4 \cdot \left[E(\xi^2)\right] \cdot E(\eta^2) \leq 0 \implies \left|E(\xi\eta)\right|^2 \leq E(\xi^2) \cdot E(\eta^2)$$

$$f(t) = E\left[(t\xi - \eta)^2\right] = 0 \text{ 有一个重根 } t = c \iff \left[2E(\xi\eta)\right]^2 - 4 \cdot \left[E(\xi^2)\right] \cdot E(\eta^2) = 0$$

$$\implies E\left[(c\xi - \eta)^2\right] = 0 \iff P\{\eta = c\xi\} = 1$$

三、相关系数

- 相关系数的性质.

▶ **性质 1** 设 ρ 为随机变量 ξ 与 η 的相关系数, 则 $|\rho| \leq 1$.

$$\rho = 1 \iff P \left\{ \frac{\xi - E(\xi)}{\sqrt{D(\xi)}} = \frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right\} = 1; \quad \rho = -1 \iff P \left\{ \frac{\xi - E(\xi)}{\sqrt{D(\xi)}} = -\frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right\} = 1$$

对随机变量 $\frac{\xi - E(\xi)}{\sqrt{D(\xi)}}$ 与 $\frac{\eta - E(\eta)}{\sqrt{D(\eta)}}$ 应用 Cauchy-Schwarz 不等式

$$\begin{aligned} & \left| E \left(\frac{\xi - E(\xi)}{\sqrt{D(\xi)}} \cdot \frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right) \right|^2 \leq E \left[\left(\frac{\xi - E(\xi)}{\sqrt{D(\xi)}} \right)^2 \right] \cdot E \left[\left(\frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right)^2 \right] \\ & \quad \quad \quad \parallel \quad \quad \quad \parallel \quad \quad \quad \parallel \\ |\rho|^2 &= \left| \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)} \cdot \sqrt{D(\eta)}} \right|^2 = \left| \frac{E \left[(\xi - E(\xi)) (\eta - E(\eta)) \right]}{\sqrt{D(\xi)} \cdot \sqrt{D(\eta)}} \right|^2 \leq \frac{E \left[(\xi - E(\xi))^2 \right]}{D(\xi)} \cdot \frac{E \left[(\eta - E(\eta))^2 \right]}{D(\eta)} = 1 \end{aligned}$$

定理: (Cauchy-Schwarz 不等式)

对任意随机变量 ξ 与 η , 如果 $E(\xi^2) < +\infty$, $E(\eta^2) < +\infty$, 则有

$$|E(\xi\eta)|^2 \leq E(\xi^2) \cdot E(\eta^2)$$

其中等式成立当且仅当存在某一常数 c 使得 $P\{\eta = c\xi\} = 1$ 成立.

三、相关系数

- 相关系数的性质.

▶ **性质 1** 设 ρ 为随机变量 ξ 与 η 的相关系数, 则 $|\rho| \leq 1$.

$$\begin{array}{l} \text{完全正相关} \\ \rho = 1 \iff P \left\{ \frac{\xi - E(\xi)}{\sqrt{D(\xi)}} = \frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right\} = 1; \quad \text{完全负相关} \\ \rho = -1 \iff P \left\{ \frac{\xi - E(\xi)}{\sqrt{D(\xi)}} = -\frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right\} = 1 \end{array}$$

$$\xrightarrow{c > 0} \rho = 1 \iff |\rho|^2 = 1 \iff \exists c \in \mathbb{R} : P\{\eta = c\xi\} = 1 \iff P \left\{ \frac{\xi - E(\xi)}{\sqrt{D(\xi)}} = \frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right\} = 1$$

$$\xrightarrow{c < 0} \rho = -1 \iff |\rho|^2 = 1 \iff \exists c \in \mathbb{R} : P\{\eta = c\xi\} = 1 \iff P \left\{ \frac{\xi - E(\xi)}{\sqrt{D(\xi)}} = -\frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right\} = 1$$

定理: (Cauchy-Schwarz 不等式)

对任意随机变量 ξ 与 η , 如果 $E(\xi^2) < +\infty$, $E(\eta^2) < +\infty$, 则有

$$|E(\xi\eta)|^2 \leq E(\xi^2) \cdot E(\eta^2)$$

其中等式成立当且仅当存在某一常数 c 使得 $P\{\eta = c\xi\} = 1$ 成立.

三、相关系数

- 相关系数的性质.

定义: 若随机变量 ξ 与 η 的相关系数 $\rho = 0$, 则称 ξ 与 η (线性) 不相关.

► **性质 2** 对随机变量 ξ 与 η , 以下事实等价:

① $\text{cov}(\xi, \eta) = 0$.

② ξ 与 η 不相关.

③ $E(\xi \cdot \eta) = E(\xi) \cdot E(\eta)$.

④ $D(\xi + \eta) = D(\xi) + D(\eta)$.

► **性质 3** 若随机变量 ξ 与 η 独立, 则 ξ 与 η 不相关.

三、相关系数

- 相关系数的性质.
- 例: 设 $\theta \sim U[0, 2\pi]$, $\xi = \cos \theta$, $\eta = \cos(\theta + a)$, a 为常数. 确定 ξ 与 η 的独立性与相关性.

$$E(\xi) = \int_0^{2\pi} \cos t \cdot \frac{1}{2\pi} dt = 0$$

$$E(\xi^2) = \int_0^{2\pi} \cos^2 t \cdot \frac{1}{2\pi} dt = \frac{1}{2}$$

$$E(\eta) = \int_0^{2\pi} \cos(t + a) \cdot \frac{1}{2\pi} dt = 0$$

$$E(\eta^2) = \int_0^{2\pi} \cos^2(t + a) \cdot \frac{1}{2\pi} dt = \frac{1}{2}$$

$$E(\xi\eta) = \int_0^{2\pi} \cos t \cdot \cos(t + a) \cdot \frac{1}{2\pi} dt = \frac{1}{2} \cos a$$

$$\rho = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)} \sqrt{D(\eta)}} = \frac{E(\xi\eta) - E(\xi) \cdot E(\eta)}{\sqrt{E(\xi^2) - [E(\xi)]^2} \sqrt{E(\eta^2) - [E(\eta)]^2}} = \frac{\frac{1}{2} \cos a - 0 \times 0}{\sqrt{\frac{1}{2} - 0^2} \sqrt{\frac{1}{2} - 0^2}} = \cos a$$

三、相关系数

- 相关系数的性质.

- 例: 设 $\theta \sim U[0, 2\pi]$, $\xi = \cos \theta$, $\eta = \cos(\theta + a)$, a 为常数. 确定 ξ 与 η 的独立性与相关性.

$$\left. \begin{array}{l} \xrightarrow{a=0} \rho = 1, \quad \xi = \eta \\ \xrightarrow{a=\pi} \rho = -1, \quad \xi = -\eta \end{array} \right\} \text{存在线性关系}$$

$$\xrightarrow{a = \frac{\pi}{2} \text{ 或 } a = \frac{3}{2}\pi} \rho = 0, \quad \xi \text{ 与 } \eta \text{ 不相关}$$

$$\xi^2 + \eta^2 = 1, \quad \xi \text{ 与 } \eta \text{ 不独立}$$

$$\rho = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)} \sqrt{D(\eta)}} = \frac{E(\xi\eta) - E(\xi) \cdot E(\eta)}{\sqrt{E(\xi^2) - [E(\xi)]^2} \sqrt{E(\eta^2) - [E(\eta)]^2}} = \frac{\frac{1}{2} \cos a - 0 \times 0}{\sqrt{\frac{1}{2} - 0^2} \sqrt{\frac{1}{2} - 0^2}} = \cos a$$

三、相关系数

- 相关系数的性质.
 - ▶ 相关系数 $\rho_{\xi\eta}$ 是随机变量 ξ 与 η 之间线性关系的一种度量.
 - ▶ 相关系数 $|\rho_{\xi\eta}| \rightarrow 1$, ξ 与 η 的线性关系越显著.
 - ▶ 相关系数 $|\rho_{\xi\eta}| \rightarrow 0$, ξ 与 η 的线性关系越不显著.
 - ▶ 即使两个随机变量 ξ 与 η 不相关, 即 $\rho_{\xi\eta} = 0$, 它们之间也可能存在其它关系.
 - ▶ 不相关性是就线性关系而言的, 而独立性是就一般关系而言的.

三、相关系数

- 相关系数的性质.

▶ **性质 4** 若随机变量 X 与 Y 服从二元正态分布, 则 X 与 Y 的独立性和不相关性是等价的.

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$

$\xrightarrow{\rho=0}$

$$= \frac{1}{2\pi\sigma_1\sigma_2} \exp \left\{ -\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2} \right\}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left\{ -\frac{(x-\mu_1)^2}{2\sigma_1^2} \right\} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left\{ -\frac{(y-\mu_2)^2}{2\sigma_2^2} \right\}$$

三、相关系数

- 相关系数的性质.

ξ	a	c	η	b	d
$P\{\xi = x_i\}$	p	$1-p$	$P\{\eta = y_j\}$	q	$1-q$

- ▶ **性质 5** 若随机变量 ξ 与 η 都是二值变量, 则其不相关性与独立性等价.

我们来证明 $\rho_{\xi\eta} = 0 \implies \xi$ 与 η 相互独立.

$$A \triangleq \{\xi = a\}, \quad \bar{A} = \{\xi = c\}$$

$$B \triangleq \{\eta = b\}, \quad \bar{B} = \{\eta = d\}$$

A 的示性函数

$$1_A \triangleq \frac{\xi - c}{a - c} = \begin{cases} 1, & \xi = a \\ 0, & \xi = c \end{cases}$$

$$1_B \triangleq \frac{\eta - d}{b - d} = \begin{cases} 1, & \eta = b \\ 0, & \eta = d \end{cases}$$

B 的示性函数

$$\text{COV}(1_A, 1_B) = \text{COV}\left(\frac{\xi - c}{a - c}, \frac{\eta - d}{b - d}\right) = \frac{1}{(a - c)(b - d)} \cdot \text{COV}(\xi, \eta) = 0$$

$$= E(1_A \cdot 1_B) - E(1_A) \cdot E(1_B)$$

$$= P\{\xi = a, \eta = b\} - P\{\xi = a\} \cdot P\{\eta = b\} = P(AB) - P(A) \cdot P(B)$$

$$\implies P\{\xi = a, \eta = b\} = P\{\xi = a\} \cdot P\{\eta = b\} \implies \{\xi = a\} \text{ 与 } \{\eta = b\} \text{ 相互独立}$$

三、相关系数

- 相关系数的性质.

ξ	a	c	η	b	d
$P\{\xi = x_i\}$	p	$1-p$	$P\{\eta = y_j\}$	q	$1-q$

- ▶ **性质 5** 若随机变量 ξ 与 η 都是二值变量, 则其不相关性与独立性等价.

我们来证明 $\rho_{\xi\eta} = 0 \implies \xi$ 与 η 相互独立.

$$A \triangleq \{\xi = a\}, \quad \bar{A} = \{\xi = c\}$$

$$B \triangleq \{\eta = b\}, \quad \bar{B} = \{\eta = d\}$$

事件独立性的性质 $\implies \{\xi = a\}$ 与 $\{\eta = d\}$ 相互独立

$\{\xi = c\}$ 与 $\{\eta = b\}$ 相互独立

$\{\xi = c\}$ 与 $\{\eta = d\}$ 相互独立

$\implies \xi$ 与 η 相互独立

$\implies P\{\xi = a, \eta = b\} = P\{\xi = a\} \cdot P\{\eta = b\} \implies \{\xi = a\}$ 与 $\{\eta = b\}$ 相互独立

三、相关系数

- 相关系数的性质.

- ▶ **性质 5** 若随机变量 ξ 与 η 都是二值变量, 则其不相关性与独立性等价.

- ◆ **推论 1** 对事件 A 与 B , 定义 A 与 B 的相关系数为

$$\rho_{AB} = \rho_{1_A 1_B} = \frac{P(AB) - P(A) \cdot P(B)}{\sqrt{P(A) \cdot P(\bar{A}) \cdot P(B) \cdot P(\bar{B})}}$$

则 A 与 B 独立的充分必要条件为 $\rho_{AB} = 0$.

- ◆ **推论 2** $|P(AB) - P(A) \cdot P(B)| \leq \frac{1}{4}$.