

概 率 论

Probability

肖磊, 2024年11月19日

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

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

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

已学知识点

● 第四章 数字特征与特征函数

▶ 数学期望

○ 定义: $E(\xi) = \begin{cases} \sum_i x_i \cdot p_i, & \text{离散型} \\ \int_{-\infty}^{\infty} x \cdot p_{\xi}(x) dx, & \text{连续型} \end{cases}$, 绝对收敛.

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

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

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: $D(\xi) = E(\xi^2) - [E(\xi)]^2$

▶ **Chebyshev 不等式**:

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2} \iff P\left\{\left|\xi - E(\xi)\right| < \varepsilon\right\} \geq 1 - \frac{D(\xi)}{\varepsilon^2}$$

▶ **协方差**:

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

▶ **和的方差**: $D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y)$

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 $D(\xi \pm \eta) = D(\xi) + D(\eta)$.

协方差的性质:

- ① $\text{cov}(X, X) = D(X)$.
- ② $\text{cov}(X, Y) = \text{cov}(Y, X)$
- ③ $\text{cov}(aX, bY) = ab \cdot \text{cov}(X, Y)$, 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 $\text{cov}(X, C) = 0$.

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

▶ **相关系数:** $\rho_{\xi\eta} \triangleq \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$

▶ **Cauchy-Schwarz 不等式:** $|E(\xi\eta)|^2 \leq E(\xi^2) \cdot E(\eta^2)$

相关系数的性质:

- ① $|\rho_{\xi\eta}| \leq 1.$
- ② 等价命题: $\text{cov}(\xi, \eta) = 0$; ξ 与 η 不相关;
 $E(\xi \cdot \eta) = E(\xi) \cdot E(\eta);$
 $D(\xi + \eta) = D(\xi) + D(\eta).$
- ③ 若随机变量 ξ 与 η 独立, 则 ξ 与 η 不相关.
- ④ 二元正态分布的独立性和不相关性是等价的.
- ⑤ 二值随机变量的不相关性与独立性等价.

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

四、矩

- 数学期望、方差、协方差、相关系数是随机变量常用的数字特征, 它们都是某种矩.
- 矩是最广泛使用的一种数字特征, 在概率论和数理统计种占有重要地位.
- 常用的矩有两种: **原点矩**和**中心矩**.

定义: 对正整数 k , 称

$$m_k = E(\xi^k), \quad k = 1, 2, 3, \dots$$

为随机变量 ξ 的 k 阶**原点矩**. 数学期望是一阶原点矩.

定义: 对正整数 k , 称

$$c_k = E\left\{\left[\xi - E(\xi)\right]^k\right\}, \quad k = 2, 3, \dots$$

为随机变量 ξ 的 k 阶**中心矩**. 方差是二阶中心矩.

四、矩

定理： 中心矩和原点矩可以相互表示.

$$c_k = E \left\{ \left[\xi - E(\xi) \right]^k \right\} = E \left\{ \sum_{i=0}^k \binom{k}{i} \xi^i \left[-E(\xi) \right]^{k-i} \right\} = \sum_{i=0}^k \binom{k}{i} E(\xi^i) \left[-E(\xi) \right]^{k-i} = \sum_{i=0}^k \binom{k}{i} m_i (-m_1)^{k-i}$$

⇒ 中心矩 c_k 可以通过原点矩 m_k 表示.

$$\begin{aligned}
 m_k = E(\xi^k) &= E \left\{ \left[(\xi - m_1) + m_1 \right]^k \right\} = E \left\{ \sum_{i=0}^k \binom{k}{i} m_1^i (\xi - m_1)^{k-i} \right\} \\
 &= \sum_{i=0}^k \binom{k}{i} m_1^i E \left[(\xi - m_1)^{k-i} \right] = \sum_{i=0}^k \binom{k}{i} c_{k-i} m_1^i
 \end{aligned}$$

⇒ 已知数学期望之后, 原点矩 m_k 也可以通过中心矩 c_k 表示.

四、矩

- 对正数 k , 还可以定义 k 阶原点绝对矩 $E\left(|\xi|^k\right)$, k 阶中心绝对矩 $E\left(|\xi - E(\xi)|^k\right)$.

原点绝对矩和中心绝对矩使用较少.

- 对于随机向量, 可以定义各种混合矩, 例如

$$E\left\{\left[\xi - E(\xi)\right]^k\left[\eta - E(\eta)\right]^l\right\}$$

称为 $k + l$ 阶混合中心矩. 协方差是二阶混合中心矩.

四、矩

- 例: 设 $\xi \sim N(0, \sigma^2)$, 求 ξ 的 k 阶原点矩和 k 阶中心矩.

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad -\infty < x < +\infty \implies E(\xi) = 0$$

$$\implies c_k = E\left[\left(\xi - E(\xi)\right)^k\right] = E(\xi^k) = m_k = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} x^k \cdot e^{-\frac{x^2}{2\sigma^2}} dx = \begin{cases} 0, & k \text{ 为奇数} \\ \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} x^k \cdot e^{-\frac{x^2}{2\sigma^2}} dx, & k \text{ 为偶数} \end{cases}$$

Γ 函数: $\Gamma(r) = \int_0^{+\infty} x^{r-1} e^{-x} dx, \quad r \in \mathbb{R}^+$

Γ 函数的性质:

- (1) $\Gamma(1) = 1$
- (2) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- (3) $\forall r \in \mathbb{R}^+, \Gamma(r+1) = r \Gamma(r)$
- (4) $\forall r \in \mathbb{N}^+, \Gamma(r+1) = r!$
- (5) $\binom{n+r-1}{n} = \frac{\Gamma(n+r)}{\Gamma(n)\Gamma(r)}$
- (6) $\frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} = \int_0^1 u^{r-1}(1-u)^{s-1} du$

$$= 2 \frac{1}{\sqrt{2\pi} \sigma} \int_0^{+\infty} x^k \cdot e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sqrt{2\pi} \sigma} \int_0^{+\infty} \left(\sqrt{2z} \sigma\right)^k \cdot e^{-z} \cdot \frac{1}{\sqrt{2}} z^{-\frac{1}{2}} \sigma dz$$

$$z = \frac{x^2}{2\sigma^2}, \quad x = \sqrt{2z} \sigma, \quad dx = \frac{1}{\sqrt{2}} z^{-\frac{1}{2}} \sigma dz$$

$$= \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \int_0^{+\infty} z^{\frac{k+1}{2}-1} e^{-z} dz = \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \cdot \Gamma\left(\frac{k+1}{2}\right)$$

四、矩

- 例: 设 $\xi \sim N(0, \sigma^2)$, 求 ξ 的 k 阶原点矩和 k 阶中心矩.

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad -\infty < x < +\infty \implies E(\xi) = 0$$

$$\implies c_k = E\left[\left(\xi - E(\xi)\right)^k\right] = E(\xi^k) = m_k = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} x^k \cdot e^{-\frac{x^2}{2\sigma^2}} dx = \begin{cases} 0, & k \text{ 为奇数} \\ \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} x^k \cdot e^{-\frac{x^2}{2\sigma^2}} dx, & k \text{ 为偶数} \end{cases}$$

Γ 函数: $\Gamma(r) = \int_0^{+\infty} x^{r-1} e^{-x} dx, \quad r \in \mathbb{R}^+$

Γ 函数的性质:

- (1) $\Gamma(1) = 1$
- (2) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- (3) $\forall r \in \mathbb{R}^+, \Gamma(r+1) = r \Gamma(r)$
- (4) $\forall r \in \mathbb{N}^+, \Gamma(r+1) = r!$
- (5) $\binom{n+r-1}{n} = \frac{\Gamma(n+r)}{\Gamma(n)\Gamma(r)}$
- (6) $\frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} = \int_0^1 u^{r-1}(1-u)^{s-1} du$

$$= \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \cdot \frac{k-1}{2} \cdot \frac{k-3}{2} \cdots \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = \sigma^k (k-1)(k-3) \cdots 3 \cdot 1$$

$$= \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \cdot \Gamma\left(\frac{k-1}{2} + 1\right) = \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \cdot \frac{k-1}{2} \cdot \Gamma\left(\frac{k-1}{2}\right)$$

$$= \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \int_0^{+\infty} z^{\frac{k+1}{2}-1} e^{-z} dz = \sqrt{\frac{2}{\pi}} \sigma^k \cdot 2^{\frac{k-1}{2}} \cdot \Gamma\left(\frac{k+1}{2}\right)$$

四、矩

- 例: 设 $\xi \sim N(0, \sigma^2)$, 求 ξ 的 k 阶原点矩和 k 阶中心矩.

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad -\infty < x < +\infty \implies E(\xi) = 0$$

$$\implies c_k = m_k = \begin{cases} 0, & k \text{ 为奇数} \\ \sigma^k (k-1)(k-3) \cdots 3 \cdot 1, & k \text{ 为偶数} \end{cases}$$

- ▶ 推广: 设 $\xi \sim N(\mu, \sigma^2)$, 则

$$\xrightarrow{k \text{ 阶中心矩}} c_k = \begin{cases} 0, & k \text{ 为奇数} \\ \sigma^k (k-1)(k-3) \cdots 3 \cdot 1, & k \text{ 为偶数} \end{cases}$$

$$\xrightarrow{k \text{ 阶原点矩}} m_k = E(\xi^k) = E\left\{ \left[(\xi - m_1) + m_1 \right]^k \right\} = \sum_{i=0}^k \binom{k}{i} c_{k-i} \cdot \mu^i$$

五、条件数学期望

离散型随机变量的条件期望

给定 (ξ, η) 的联合分布律, 可求得 ξ, η 的边际分布律,

随机变量 $(\eta | \xi = x_i)$ 的分布律为

$(\eta \xi = x_i)$	y_1	y_2	...	y_j	...
$P\{\eta = y_j \xi = x_i\}$	$\frac{p(x_i, y_1)}{p_1(x_i)}$	$\frac{p(x_i, y_2)}{p_1(x_i)}$...	$\frac{p(x_i, y_j)}{p_1(x_i)}$...

定义 $\xi = x_i$ 条件下 η 的条件数学期望为

$$E(\eta | \xi = x_i) = \sum_j y_j \cdot P\{\eta = y_j | \xi = x_i\}$$

$$= \sum_j y_j \cdot \frac{p(x_i, y_j)}{p_1(x_i)}$$

- $E(\eta | \xi)$ 反映了随机变量 η 的平均值对随机变量 ξ 的依赖.
- η 关于 ξ 的条件期望 $E(\eta | \xi)$ 是一个随机变量, 它取值为 $E(\eta | \xi = x_i)$ 的概率为 $P\{\xi = x_i\}$.

	η	y_1	y_2	...	y_j	...	$p_1(\cdot)$
ξ		y_1	y_2	...	y_j	...	$p_1(\cdot)$
x_1		$p(x_1, y_1)$	$p(x_1, y_2)$...	$p(x_1, y_j)$...	$p_1(x_1)$
x_2		$p(x_2, y_1)$	$p(x_2, y_2)$...	$p(x_2, y_j)$...	$p_1(x_2)$
\vdots		\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
x_i		$p(x_i, y_1)$	$p(x_i, y_2)$...	$p(x_i, y_j)$...	$p_1(x_i)$
\vdots		\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
$p_2(\cdot)$		$p_2(y_1)$	$p_2(y_2)$...	$p_2(y_j)$...	1

五、条件数学期望

离散型随机变量的条件期望

给定 (ξ, η) 的联合分布律, 可求得 ξ, η 的边际分布律,

随机变量 $(\xi | \eta = y_j)$ 的分布律为

$(\xi \eta = y_j)$	x_1	x_2	...	x_i	...
$P\{\xi = x_i \eta = y_j\}$	$\frac{p(x_1, y_j)}{p_2(y_j)}$	$\frac{p(x_2, y_j)}{p_2(y_j)}$...	$\frac{p(x_i, y_j)}{p_2(y_j)}$...

定义 $\eta = y_j$ 条件下 ξ 的条件数学期望为

$$E(\xi | \eta = y_j) = \sum_i x_i \cdot P\{\xi = x_i | \eta = y_j\}$$

$$= \sum_i x_i \cdot \frac{p(x_i, y_j)}{p_2(y_j)}$$

- $E(\xi | \eta)$ 反映了随机变量 ξ 的平均值对随机变量 η 的依赖.
- ξ 关于 η 的条件期望 $E(\xi | \eta)$ 是一个随机变量, 它取值为 $E(\xi | \eta = y_j)$ 的概率为 $P\{\eta = y_j\}$.

η	y_1	y_2	...	y_j	...	$p_1(\cdot)$
ξ						
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$...	$p(x_1, y_j)$...	$p_1(x_1)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$...	$p(x_2, y_j)$...	$p_1(x_2)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$...	$p(x_i, y_j)$...	$p_1(x_i)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
$p_2(\cdot)$	$p_2(y_1)$	$p_2(y_2)$...	$p_2(y_j)$...	1

五、条件数学期望

- 离散型随机变量的条件期望

- 例: 设 ξ, η 的联合分布律为

可求得 ξ, η 的边际分布律为

$\xi \backslash \eta$	-1	0	2	$p_{\xi}(x)$
0	0.1	0.2	0	0.3
1	0.3	0.05	0.1	0.45
2	0.15	0	0.1	0.25
$p_{\eta}(y)$	0.55	0.25	0.2	

$$\Rightarrow E(\eta | \xi = 0) = \sum_j y_j \cdot \frac{p(0, y_j)}{p_{\xi}(0)} = -1 \times \frac{0.1}{0.3} + 0 \times \frac{0.2}{0.3} + 2 \times \frac{0}{0.3} = -\frac{1}{3}$$

$$E(\eta | \xi = 1) = \sum_j y_j \cdot \frac{p(1, y_j)}{p_{\xi}(1)} = -1 \times \frac{0.3}{0.45} + 0 \times \frac{0.05}{0.45} + 2 \times \frac{0.1}{0.45} = -\frac{2}{9}$$

$$E(\eta | \xi = 2) = \sum_j y_j \cdot \frac{p(2, y_j)}{p_{\xi}(2)} = -1 \times \frac{0.15}{0.25} + 0 \times \frac{0}{0.25} + 2 \times \frac{0.1}{0.25} = \frac{1}{5}$$

$\Rightarrow \eta$ 关于 ξ 的条件期望 $E(\eta | \xi)$ 是一个随机变量, 其分布律为

$E(\eta \xi)$	$-\frac{1}{3}$	$-\frac{2}{9}$	$\frac{1}{5}$
p_i	0.3	0.45	0.25

五、条件数学期望

- 离散型随机变量的条件期望

- 例: 设 ξ, η 的联合分布律为

可求得 ξ, η 的边际分布律为

$\xi \backslash \eta$	-1	0	2	$p_\xi(x)$
0	0.1	0.2	0	0.3
1	0.3	0.05	0.1	0.45
2	0.15	0	0.1	0.25
$p_\eta(y)$	0.55	0.25	0.2	

$$\Rightarrow E(\xi | \eta = -1) = \sum_i x_i \cdot \frac{p(x_i, -1)}{p_\eta(-1)} = 0 \times \frac{0.1}{0.55} + 1 \times \frac{0.3}{0.55} + 2 \times \frac{0.15}{0.55} = \frac{12}{11}$$

$$E(\xi | \eta = 0) = \sum_i x_i \cdot \frac{p(x_i, 0)}{p_\eta(0)} = 0 \times \frac{0.2}{0.25} + 1 \times \frac{0.05}{0.25} + 2 \times \frac{0}{0.25} = \frac{1}{5}$$

$$E(\xi | \eta = 2) = \sum_i x_i \cdot \frac{p(x_i, 2)}{p_\eta(2)} = 0 \times \frac{0}{0.2} + 1 \times \frac{0.1}{0.2} + 2 \times \frac{0.1}{0.2} = \frac{3}{2}$$

$\Rightarrow \xi$ 关于 η 的条件期望 $E(\xi | \eta)$ 是一个随机变量, 其分布律为

$E(\xi \eta)$	$\frac{12}{11}$	$\frac{1}{5}$	$\frac{3}{2}$
p_i	0.55	0.25	0.2

五、条件数学期望

- 连续型随机变量的条件期望

定义: 在 $\xi = x$ 的条件下, η 的条件数学期望定义为

$$E(\eta | \xi = x) = \int_{-\infty}^{+\infty} y \cdot p(y | x) dy$$

- ▶ 条件数学期望具有数学期望的所有性质.
- ▶ η 关于 ξ 的条件数学期望 $E(\eta | \xi)$ 是一个随机变量, 在 $E(\eta | \xi = x)$ 处的密度函数为 $p_{\xi}(x)$.

五、条件数学期望

- 连续型随机变量的条件期望

- 例: 设 ξ, η 服从二元正态分布 $(\xi, \eta) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

第三章计算条件密度得 $\rightarrow (\eta | \xi = x) \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2)\right)$

$$\Rightarrow E(\eta | \xi = x) = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1)$$

从而有随机变量 $\Rightarrow E(\eta | \xi) = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (\xi - \mu_1) \sim N(\mu_2, \rho^2 \sigma_2^2)$

五、条件数学期望

- 连续型随机变量的条件期望
→ 联合密度 $p(x, y) = p_\eta(y) \cdot p_{\xi|\eta}(x|y)$

定理: (重期望公式) 设 (ξ, η) 为二维随机向量, 且 $E(\xi)$ 存在, 则有

$$E(\xi) = E\left[E(\xi|\eta)\right]$$

▶ 就连续情形来证明.

$$\begin{aligned}
 E\left[E(\xi|\eta)\right] &= \int_{-\infty}^{+\infty} E(\xi|\eta=y) \cdot p_\eta(y) dy = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} x \cdot p_{\xi|\eta}(x|y) dx \right] \cdot p_\eta(y) dy \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot p_{\xi|\eta}(x|y) \cdot p_\eta(y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot p(x, y) dx dy \\
 &= \int_{-\infty}^{+\infty} x \left[\int_{-\infty}^{+\infty} p(x, y) dy \right] dx = \int_{-\infty}^{+\infty} x \cdot p_\xi(x) dx = E(\xi)
 \end{aligned}$$

五、条件数学期望

- 连续型随机变量的条件期望

- ▶ 重期望公式的应用:
$$\begin{cases} E(\xi) = \sum_j E(\xi | \eta = y_j) \cdot P\{\eta = y_j\} \\ E(\xi) = \int_{-\infty}^{+\infty} E(\xi | \eta = y) \cdot p_\eta(y) dy \end{cases}$$

- ▶ 该公式提供了一个在大范围求平均的一种思想方法, 即所谓的**两次平均法**.

4.5 特征函数

- 一、特征函数的定义
- 二、特征函数的性质
- 三、逆转公式与唯一性定理
- 四、分布函数的再生性
- 五、多元特征函数

一、特征函数的定义

- 分布函数及其密度无疑是描述随机变量概率规律的有力工具，可以方便地解决许多与随机变量有关的概率问题.
- 分布函数的不足：
 - ▶ 分布函数本身的分析性质不太好，它只是一个单边连续的有界非降函数.
 - ▶ 相互独立的随机变量之和的分布函数等于各分布函数的卷积，这在计算上带来不少麻烦.
- 另一方面，数字特征也只反映了概率分布的某些侧面，一般并不能通过它来确定分布函数.
- **特征函数**：即能完全决定分布函数，又具有良好的分析性质.

一、特征函数的定义

定义： 如果 ξ 与 η 都是概率空间 (Ω, \mathcal{F}, P) 上的实值随机变量，则称 $\zeta = \xi + i\eta$ 为复随机变量.

- ▶ 对复随机变量的研究本质上是对二维随机变量的研究.
- ▶ 如果二维随机变量 (ξ_1, η_1) 与 (ξ_2, η_2) 相互独立，则称复随机变量 $\zeta_1 = \xi_1 + i\eta_1$ 与 $\zeta_2 = \xi_2 + i\eta_2$ 相互独立.
- ▶ 定义复随机变量 $\zeta = \xi + i\eta$ 的数学期望为

$$E(\zeta) = E(\xi) + iE(\eta)$$

一、特征函数的定义

- 对于复随机变量，可平行的定义或建立一系列结果. 例如：

- ▶ 若 $\zeta_1, \zeta_2, \dots, \zeta_n$ 是相互独立的复随机变量，则

$$E(\zeta_1 \zeta_2 \cdots \zeta_n) = E(\zeta_1) E(\zeta_2) \cdots E(\zeta_n)$$

- ▶ 若 $g(x)$ 是一元 Borel 可测函数， $\eta = g(\xi)$ ，则

$$E(e^{it\eta}) = E(e^{it \cdot g(\xi)}) = \int_{-\infty}^{+\infty} e^{it \cdot g(x)} dF_{\xi}(x) = \begin{cases} \sum_k e^{it \cdot g(x_k)} \cdot p_k, & \text{离散型} \\ \int_{-\infty}^{+\infty} e^{it \cdot g(x)} p_{\xi}(x) dx, & \text{连续型} \end{cases}$$

- ▶ Euler 公式： $e^{ix} = \cos x + i \sin x$.

一、特征函数的定义

定义：如果随机变量 ξ 的分布函数为 $F_\xi(x)$ ，则称

$$f_\xi(t) = E(e^{it\xi}) = \begin{cases} \sum_k e^{it \cdot x_k} \cdot p_k, & \text{离散型} \\ \int_{-\infty}^{+\infty} e^{it \cdot x} p_\xi(x) dx, & \text{连续型} \end{cases}$$

为 ξ 的特征函数.

- ▶ 特征函数是一个实变量的复值函数，由于 $|e^{itx}| = 1$ ，所以它对一切实数 t 都有意义.
- ▶ 显然特征函数只与分布函数有关，因此又称**某一分布函数的特征函数**.