

基础概率论 - 李贤平 - 笔记

作者: 若水

邮箱: ethanmxzhou@163.com 主页: helloethanzhou.github.io

时间: July 18, 2024



致谢

感谢 勇敢的 自己

目录

第一章	事件与概率	1
1.1	随机现象与统计规律性	1
1.2	样本空间与事件	1
1.3	古典概型	2
1.4	几何概率	2
1.5	概率空间	2
	条件概率与统计独立性	4
	条件概率,全概率公式, Bayes 公式	4
	事件独立性	
	Bernoulli 试验与直线上的随机游动	5
2.4	二项分布与 Poisson 分布	6
笋=音	随机变量与分布函数	8
	随机变量与其分布	8
3.1	3.1.1 随机变量	8
	3.1.2 离散型随机变量	8
	3.1.3 连续型随机变量	9
3.2	随机向量,随机变量的独立性	9
3.2	3.2.1 随机向量及其分布	9
	3.2.2 边际分布	11
	3.2.3 条件分布	12
	3.2.4 随机变量的独立性	12
3 3	随机变量的函数及其分布	13
3.3	3.3.1 Borel 函数与随机变量的函数	13
	3.3.2 单个随机变量的函数的分布律	
		13
	3.3.3 多 随机文里的函数的分型件	14
第四章	数字特征与特征函数	16
4.1	数学期望	16
4.2	方差,相关系数,矩	17
	4.2.1 方差	17
	4.2.2 协方差	18
	4.2.3 相关系数与相关性	18
	4.2.4 矩	19
	4.2.5 分位数	20
	4.2.6 条件数学期望与最小二乘回归	20
4.3	母函数	21
4.4	特征函数	21
	4.4.1 特征函数	21
	4.4.2 特征函数的唯一性	22
第五章	: 极限定理	23

		目录
5.1	Bernoulli 试验场合的极限定理	23
	5.1.1 大数定律与中心极限定理	23
	5.1.2 大数定律	23
	5.1.3 De Moivre—Laplace 极限定理	24
5.2	收敛性	25
	5.2.1 分布函数弱收敛	25
	5.2.2 Lévy—Cramer 定理	25
	5.2.3 随机变量的收敛性	26
	5.2.4 Bochner—Khinchin 定理	
5.3	独立同分布场合的极限定理	27
	5.3.1 Khinchin 大数定律	27
	5.3.2 中心极限定理	27
5.4	强大数定律	27
	5.4.1 Borel 强大数定律	27
	5.4.2 Kolmogorov 强大数定律	
5.5	中心极限定理	29
	收敛性,大数定律与中心极限定理	31
A.1	收敛性	
	A.1.1 分布函数	
	A.1.2 随机变量的四大收敛性	
A.2	大数定律	
	A.2.1 大数定理的定义	
	A.2.2 经典大数定律	
A.3	强大数定律	
	A.3.1 强大数定律的定义	33
	A.3.2 经典强大数定律	34
A.4	中心极限定理	
	A.4.1 中心极限定理的定义	
	A.4.2 经典中心极限定理	35
附录 B	概率模型	37

第一章 事件与概率

1.1 随机现象与统计规律性

必然事件: 在一定条件下, 必然会发生的事情。

不可能事件: 在一定条件下,不然不会发生的事情。

随机现象: 在基本条件不变的情况下,一系列试验或观察会得到不同的结果的现象。

随机事件: 随机现象出现的结果。

频率: 对于随机事件 A,若在 N 次试验中出现了 n 次,则称 $F_N(A) = n/N$ 为随机事件 A 在 N 次试验中出现的频率。

概率:对于随机事件 A,表示该事件发生的可能性大小的数 P(A) 称为该事件的概率。

命题 1.1.1 (频率的性质)

- 1. $F_N(A) \ge 0$
- 2. 若记必然事件为 Ω , 则有 $F_N(\Omega) = 1$ 。
- 3. 若 $A \cap B = \emptyset$, 则 $F_N(A \cup B) = F_N(A) + F_N(B)$

命题 1.1.2 (概率的性质)

- 1. 非负性: $P(A) \ge 0$
- 2. 规范性: $P(\Omega) = 1$

$$P\left(\sum_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$

1.2 样本空间与事件

样本点:随机试验可能出现的结果称为样本点,记作 ω 。

样本空间:样本点全体构成样本空间,记作 Ω 。

样本空间的类型:

- 1. 有限个样本点
- 2. 无穷可列个样本点
- 3. 无穷不可列个样本点
- 4. 三维空间
- 5. 函数空间

事件:某些样本点构成的集合。

1.3 古典概型

命题 1.3.1 (模型与计算公式)

- 2. 若 $A = \omega_{k_1} + \cdots + \omega_{k_m}$, 则 $P(A) = \frac{m}{n}$ o

命题 1.3.2 (组合公式)

- 1. 排列
 - (a). 有放回的选取并排列: n^r
 - (b). 不放回的选取并排列: $A_n^r = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}$
 - (c). 全排列: n!
 - (d). 圆排列: (n-1)!
- 2. 组合
 - (a). 不放回的选取: $C_n^r = \binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!}$
 - (b). 分为多个部分: $\frac{n!}{r_1!\cdots r_k!}$ (c). 有放回的选取: $\binom{n+r-1}{r}$

 - (d). 全错排: $a_n = n!(\frac{1}{0!} \frac{1}{1!} + \frac{1}{2!} \ldots + (-1)^n \frac{1}{n!})$
- 3. 组合公式
 - (a). $\binom{n}{k} = \binom{n}{n-k}$
 - (b). $\binom{n}{0} + \dots + \binom{n}{n} = 2^n$
 - (c). $\binom{a}{0}\binom{b}{n} + \cdots + \binom{a}{n}\binom{b}{0} = \binom{a+b}{n}$
 - (d). $\binom{-\alpha}{k} = (-1)^k \binom{\alpha+k-1}{k}$

1.4 几何概率

计算公式: $P(A) = m(A)/m(\Omega)$

命题 1.4.1 (几何概率的性质)

- 1. 非负性: $P(A) \ge 0$
- 2. 规范性: $P(\Omega) = 1$
- 3. 可数可加性: 若 A_1, A_2, \cdots 两两不相容,则

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

1.5 概率空间

定义 1.5.1 (σ -域)

称集族 $\Sigma \subset \mathcal{P}(\Omega)$ 为样本空间 Ω 上的 σ -域,如果成立

- 1. $\emptyset \in \Sigma$
- 2. 如果 $A \in \Sigma$, 那么 $\Omega \setminus A \in \Sigma$;

3. 如果 $\{A_n\}_{n=1}^{\infty} \subset \Sigma$, 那么

$$\bigcup_{n=1}^{\infty} A_n \in \Sigma$$

定义 1.5.2 (概率)

对于样本空间 Ω 上的 σ -域 $\Sigma \subset \mathcal{P}(\Omega)$, 称映射 $P: \Sigma \to [0,1]$ 为概率, 如果成立如下命题。

- 1. $P(\Omega) = 1$
- 2. 对于不相容序列 $\{A_n\}_{n=1}^{\infty}\subset \Sigma$, 成立

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

定义 1.5.3 (概率空间)



定义 1.5.4 (事件域)

称 Σ 为概率空间 (Ω, Σ, P) 的事件域, 其中的元素称为事件。



命题 1.5.1 (概率的性质)

- 1. 对于任意 $A \in \Omega$,成立 $P(A^c) = 1 P(A)$ 。
- 2. 如果 $B \subset A$, 那么 $P(A \setminus B) = P(A) P(B)$ 。
- 3. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$
- 4. Bool 不等式: $P(A \cup B) \leq P(A) + P(B)$
- 5. Bonferroni 不等式: $P(A \cap B) \ge P(A) + P(B) 1$
- 6. 下连续性:对于单调递减事件序列 $\{A_n\}_{n=1}^{\infty}$,成立

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

7. 上连续性: 对于单调递增事件序列 $\{A_n\}_{n=1}^{\infty}$, 成立

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

8.
$$P\left(\sum_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} P(A_n)$$



第二章 条件概率与统计独立性

2.1 条件概率,全概率公式, Bayes 公式

定义 2.1.1 (条件概率)

对于概率空间 (Ω, Σ, P) , 如果事件 $B \in \Sigma$ 成立 P(B) > 0, 那么对于事件 $A \in \Sigma$, 称

$$P(A \mid B) = \frac{P(AB)}{P(B)}$$

为在事件 B 发生的条件下事件 A 发生的条件概率。

命题 2.1.1 (条件概率的性质)

- 1. 非负性: $P(A \mid B) \ge 0$
- 2. 规范性: $P(\Omega \mid B) = 1$
- 3. 可数可加性: 若 A_1, A_2, \cdots 两两不相容, 则 $P\left(\sum_{k=1}^{\infty} A_k \mid B\right) = \sum_{k=1}^{\infty} P(A_k \mid B)$
- 4. 乘法原理:

$$P(AB) = P(B)P(A \mid B)$$

5. 推广的乘法公式:

$$P(A_1 \cdots A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2) \cdots P(A_n \mid A_1A_2 \cdots A_{n-1})$$

定理 2.1.1 (全概率公式)

对于样本空间 Ω 的划分

$$\Omega = \bigsqcup_{n=1}^{\infty} A_n$$

那么对于事件 B, 成立全概率公式

$$P(B) = \sum_{k=1}^{\infty} P(A_k) P(B \mid A_k)$$

定理 2.1.2 (Bayes 公式)

若事件 A_1, A_2, \cdots 两两不相容,且对于事件 $B = \sum_{k=1}^{\infty} A_k B$,存在 Bayes 公式

$$P(A_k \mid B) = \frac{P(A_k)P(B \mid A_k)}{\sum_{i=1}^{\infty} P(A_i)P(B \mid A_i)}$$

 $^{\circ}$

2.2 事件独立性

定义 2.2.1 (两个事件的独立性)

称事件A和B是统计独立的,如果

$$P(AB) = P(A)P(B)$$

定义 2.2.2 (三个事件的独立性)

称事件 A, B, C 是统计独立的, 如果

$$\begin{cases} P(AB) = P(A)P(B) \\ P(BC) = P(B)P(C) \\ P(CA) = P(C)P(A) \\ P(ABC) = P(A)P(B)P(C) \end{cases}$$

命题 2.2.1

- 1. 若事件 A, B 独立, 且 P(B) > 0, 则 $P(A \mid B) = P(A)$ 。
- 2. 若事件 A, B 独立,则事件 A, B^c 独立。

定义 2.2.3 (试验的独立性)

令 \mathscr{A}_k 为第 k 次试验有关的事件全体。称试验 $\mathscr{A}_1, \cdots, \mathscr{A}_n$ 是相互独立的,如果对于任意的 $A_k \in \mathscr{A}_k$,成立

$$P(A_1 \cdots A_n) = P(A_1) \cdots P(A_n)$$

2.3 Bernoulli 试验与直线上的随机游动

定义 2.3.1 (Bernoulli 试验)

只存在两种可能结果的试验称为 Bernoulli 试验。

定义 2.3.2 (n 重 Bernoulli 试验)

称 n 次 Bernoulli 试验为 n 重 Bernoulli 试验,如果成立如下命题。

- 1. 每次实验至多出现两个可能结果 A 与 A^c 之一。
- 2. A 在每次试验中出现的概率 p 不变。
- 3. 各次试验相互独立。

命题 2.3.1 (几何分布的无记忆性)

在 Bernoulli 试验中,已知在前n次试验中没有成功,则首次成功所在需要的时间满足几何分布。在离散型分布中,仅有几何分布满足此性质。

定理 2.3.1 (分赌注问题)

甲、乙两个赌徒中止赌博,若甲再胜n场则可赢得赌注,乙再胜m场则可赢得赌注。甲在每局获胜的概率为p,则甲赢得赌注的概率为

$$p_{\Psi} = \sum_{k=0}^{m-1} \binom{n+k-1}{k} p^n (1-p)^k$$
$$= \sum_{k=n}^{\infty} \binom{m+k-1}{k} p^k (1-p)^m$$
$$= \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}$$

定理 2.3.2 (直线上的随机游动)

考虑 x 轴上的一个质点,规定其只能位于整数点,在 t=0 时刻,位于初始位置 $x=a\in\mathbb{Z}$,以后每隔单位时间,分别以概率 p 及 1-p 向正的或负的方向移动一个单位。

1. 无限制随机游动: 若质点在 t=0 时刻从原点出发, 质点在 t=n 时刻位于 k 的概率为

$$p(n,k) = \begin{cases} \left(\frac{n}{n+k}\right) p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}}, & n \equiv k \mod 2\\ 0, & n \not\equiv k \mod 2 \end{cases}$$

2. 两端带有吸收壁的随机游动: 若质点在 t=0 时刻从 $x=n\in(a,b)$ 出发, 而在 $x=a\in\mathbb{Z}$ 和 $x=b\in\mathbb{Z}$ 处各有一个吸收壁, 质点碰到吸收壁后将不再运动, 则质点最终在 a 点被吸收的概率为

$$f_a(n) = \begin{cases} \frac{(\frac{1-p}{p})^n - (\frac{1-p}{p})^b}{(\frac{1-p}{p})^a - (\frac{1-p}{p})^b}, & p \neq \frac{1}{2} \\ \frac{n-b}{a-b}, & p = \frac{1}{2} \end{cases}$$

质点最终在 b 点被吸收的概率为

$$f_b(n) = \begin{cases} \frac{(\frac{1-p}{p})^n - (\frac{1-p}{p})^a}{(\frac{1-p}{p})^b - (\frac{1-p}{p})^a}, & p \neq \frac{1}{2} \\ \frac{n-a}{b-a}, & p = \frac{1}{2} \end{cases}$$

定义 2.3.3 (多项分布)

n 次重复独立试验且每次试验出现的可能结果为 A_1,\cdots,A_r ,其中 $P(A_k)=p_k\geq 0$ 且 $p_1+\cdots+p_r=1$,则 n 次试验中 A_k 出现 n_k 次的概率为

$$P = \frac{n!}{n_1! \cdots n_r!} p_1^{n_1} \cdots p_r^{n_r}$$

其中 $n_k \geq 0$ 且 $n_1 + \cdots + n_r = n$ 。

2.4 二项分布与 Poisson 分布

命题 2.4.1

函数 b(k; n, p) 关于 k 先递增后递减, 且

$$b_{\max}(k; n, p) = b([(n+1)p]; n, p)$$

定理 2.4.1 (Poisson 定理)

在独立实验中,以与 n 有关的常数 p_n 代表事件在实验中出现的概率。若 $np_n
ightarrow \lambda$,则当 $n
ightarrow \infty$ 时,

$$b(k; n, p) \to \frac{\lambda^k}{k!} e^{-\lambda}$$

\<u>\</u>

第三章 随机变量与分布函数

3.1 随机变量与其分布

3.1.1 随机变量

定义 3.1.1 (随机变量)

对于概率空间 (Ω, Σ, P) , 称函数 $\xi: \Omega \to \mathbb{R}$ 随机变量, 如果

$$\xi^{-1}(\mathscr{B}) \subset \Sigma$$

定义 3.1.2 (概率分布)

对于概率空间 (Ω, Σ, P) , 定义随机变量 ξ 在 Borel 集 $B \in \mathcal{B}$ 上的概率分布为 $P\{\xi^{-1}(B)\}$ 。

定义 3.1.3 (分布函数)

对于概率空间 (Ω, Σ, P) , 定义随机变量 ξ 的分布函数为

$$F(x) = P\{\xi^{-1}(-\infty, x)\}, \qquad x \in \overline{\mathbb{R}}$$

命题 3.1.1 (分布函数的性质)

- 1. $F(-\infty) = 0$, $F(+\infty) = 1$
- 2. 单调性: 若 a < b, 则 $F(a) \le F(b)$ 。
- 3. 左连续性: $F(x^{-}) = F(x)$
- 4. 分布函数至多仅有可列个不连续点。
- 5. 对于分布函数 F(x) 存在 Lebesgue 分解

$$F(x) = c_1 F_1(x) + c_2 F_2(x) + c_3 F_3(x)$$

其中 $F_1(x)$ 为跳跃函数, $F_1(x)$ 为绝对连续函数, $F_1(x)$ 为奇异函数。

定理 3.1.1 (随机变量的存在性定理)

如果 F(x) 是左连续的单调不减函数,且 $F(-\infty)=0, F(+\infty)=1$,那么存在概率空间 (Ω,Σ,P) 及其上的随机变量 ξ ,使 ξ 的分布函数为 F(x)。

3.1.2 离散型随机变量

定义 3.1.4 (概率分布)

令 $\{x_k\}_{k=1}^\infty$ 为离散型随机变量 ξ 的所有可能值, $p(x_k)$ 是 ξ 取 x_k 的概率,称 $\{p(x_k):k\in\mathbb{N}^*\}$ 为随机变量 ξ 的概率分布。

定义 3.1.5 (分布函数)

$$F(x) = \sum_{x_k < x} p(x_k)$$

3.1.3 连续型随机变量

定义 3.1.6 (分布密度函数)

$$p(x), \quad x \in \mathbb{R}$$

定义 3.1.7 (分布函数)

$$F(x) = \int_{-\infty}^{x} p(t) dt, \quad x \in \mathbb{R}$$

命题 3.1.2 (连续型随机变量的性质)

- 1. $p(x) \ge 0$
- $2. \int_{-\infty}^{\infty} p(x) dx = 1$
- 3. $P\{a \le \xi < b\} = F(b) F(a) = \int_a^b p(x) dx$
- 4. $P\{\xi = c\} = 0$

定义 3.1.8 (正态分布 $N(\mu, \sigma^2)$)

1. 密度函数:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

2. 分布函数:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \qquad x \in \mathbb{R}$$

定义 3.1.9 (标准正态分布 N(0,1))

记标准正态分布 N(0,1) 分布密度函数与分布函数分别为 $\varphi(x)$ 和 $\Phi(x)$, 那么

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \qquad x \in \mathbb{R}$$

那么对于正态分布 $N(\mu, \sigma^2)$ 的分布密度函数 p(x) 与分布函数 F(x),成立

$$p(x) = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right), \qquad F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

3.2 随机向量, 随机变量的独立性

3.2.1 随机向量及其分布

定义 3.2.1 (n 维随机向量)

对于概率空间 (Ω, Σ, P) 上的随机变量 ξ_1, \dots, ξ_n , 定义 n 维随机向量为

$$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)$$

定义 3.2.2 (联合分布函数)

对于概率空间 (Ω, Σ, P) , 称 n 元函数

$$F(x_1,\dots,x_n) = P\{\boldsymbol{\xi}^{-1}((-\infty,x_1)\times\dots\times(-\infty,x_n))\}\$$

为随机向量 $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)$ 的联合分布函数。

命题 3.2.1 (多元分布函数的性质)

- 1. 单调性:关于每个变元是单调不减函数。
- 2. $F(x_1, \dots, -\infty, \dots, x_n) = 0$ $F(+\infty, \dots, +\infty) = 1$
- 3. 左连续性:关于每个变元左连续。

定义 3.2.3 (多元分布密度函数)

对于连续型随机向量 $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$, 称 $p(x_1, \dots, x_n)$ 称为多元分布密度函数, 如果

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p(t_1, \dots, t_n) dt_1 \dots dt_n$$

定义 3.2.4 (多项分布)

$$P\{\xi_1 = k_1, \dots, \xi_r = k_r\} = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}$$

其中 $p_1 + \cdots + p_r = 1$ 且 $k_1 + \cdots + k_r = n$.

定义 3.2.5 (多元超几何分布)

$$P\{\xi_1 = n_1, \cdots, \xi_r = n_r\} = \frac{\binom{N_1}{n_1} \cdots \binom{N_r}{n_r}}{\binom{N}{n}}$$

其中 $n_1 + \cdots + n_r = n$ 。

定义 3.2.6 (均匀分布)

$$p(x_1, \dots, x_n) = \begin{cases} \frac{1}{S}, & (x_1, \dots, x_n) \in G \\ 0, & (x_1, \dots, x_n) \notin G \end{cases}$$

定义 3.2.7 (多元正态分布 $N(\mu, \Sigma)$)

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right)$$

其中 μ 为n阶向量, Σ 为n阶正定对称矩阵。

*

3.2.2 边际分布

定义 3.2.8 (边际分布)

考虑二维随机向量 (ξ,η) , 令

$$P\{\xi = x, \eta = y\} = p(x, y)$$

$$P\{\xi = x\} = p_1(x)$$

$$P\{\eta = y\} = p_2(y)$$

则

$$p_1(x) = \sum_{y \in \mathbb{R}} p(x, y), \qquad p_2(y) = \sum_{x \in \mathbb{R}} p(x, y)$$

称 $\{p_1(x):x\in\mathbb{R}\}$ 与 $\{p_2(y):y\in\mathbb{R}\}$ 为 $\{p(x,y):x,y\in\mathbb{R}\}$ 的边际分布。

定义 3.2.9 (边际分布函数)

考虑二维随机向量 (ξ,η) , 其分布函数为 F(x,y), 称

$$F_1(x) = F(x, +\infty), \qquad F_2(y) = F(+\infty, y)$$

为F(x,y)的边际分布函数。

定义 3.2.10 (边际分布密度函数)

考虑二维连续型随机向量 (ξ,η) , 其分布函数为 F(x,y), 密度函数为 p(x,y), 称

$$p_1(x) = \int_{-\infty}^{+\infty} p(x, y) dy, \qquad p_2(y) = \int_{-\infty}^{+\infty} p(x, y) dx$$

为p(x,y)的边际分布密度函数。

定义 3.2.11 (二元正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$)

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

其中 μ_1,μ_2 为两个边际分布的数学期望, σ_1,σ_2 为两个边际分布的标准差, ρ 为二元正态分布的相关系数,且构成协方差矩阵

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$$

命题 3.2.2 (二元正态分布密度函数的典型分解)

二元正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 存在如下两个分解

$$\begin{split} p(x,y) &= \frac{1}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{(y-(\mu_2+\rho\frac{\sigma_2}{\sigma_1^2}(x-\mu_1)))^2}{2\sigma_2^2(1-\rho^2)}} \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} \mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{(x-(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(x-\mu_2)))^2}{2\sigma_1^2(1-\rho^2)}} \end{split}$$

- 1. 第一式的第一部分为 $N(\mu_1,\sigma_1)$ 的密度函数,第二部分为 $N(\mu_2+\rho\frac{\sigma_1}{\sigma_2}(x-\mu_1),\sigma_2^2(1-\rho^2))$ 的密度函数。
- 2. 第二式的第一部分为 $N(\mu_2, \sigma_2)$ 的密度函数,第二部分为 $N(\mu_1 + \rho \frac{\sigma_2}{\sigma_1}(y \mu_2), \sigma_1^2(1 \rho^2))$ 的密度函数。

定义 3.2.12 (二元正态分布的边际分布)

二元正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的边际分布密度函数为

$$p_1(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \qquad p_2(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$$

因此二元正态分布的边际分布仍为正态分布。

3.2.3 条件分布

定义 3.2.13 (离散型随机变量的条件分布)

对于二维离散型随机向量 (ξ,η) , η 关于 ξ 的条件分布为

$$P{\eta = y \mid \xi = x} = \frac{p(x,y)}{p_1(x)}$$

定义 3.2.14 (连续型随机变量的条件分布)

对于二维连续型随机向量 (ξ,η) , η 关于 ξ 的条件分布为

$$P\{\eta < y \mid \xi = x\} = \lim_{\Delta x \to 0} \frac{F(x + \Delta x, y) - F(x, y)}{F(x + \Delta x, \infty) - F(x, \infty)}$$
$$= \lim_{\Delta x \to 0} \frac{\int_{x}^{x + \Delta x} \int_{-\infty}^{y} p(u, v) du dv}{\int_{x}^{x + \Delta x} \int_{-\infty}^{\infty} p(u, v) du dv}$$

若 $p_1(x) \neq 0$,则

$$p(y \mid x) = \frac{p(x, y)}{p_1(x)}$$

3.2.4 随机变量的独立性

定义 3.2.15 (离散型随机变量的独立性)

称离散型随机变量 ξ_1, \dots, ξ_n 是相互独立的,如果对于任意的 x_1, \dots, x_n ,成立

$$P\{\xi_1 = x_1, \dots, \xi_n = x_n\} = P\{\xi_1 = x_1\} \dots P\{\xi_n = x_n\}$$

定义 3.2.16 (随机变量的独立性)

称随机变量 ξ_1, \dots, ξ_n 相互独立,如果成立如下命题之一。

1. 对于任意的 x_1, \dots, x_n , 成立

$$P\{\xi_1 < x_1, \cdots, \xi_n < x_n\} = P\{\xi_1 < x_1\} \cdots P\{\xi_n < x_n\}$$

2. 对于任意的 x_1, \dots, x_n , 成立

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

推论 3.2.1

如果随机变量 ξ_1, \dots, ξ_n 相互独立, 那么对于任意的 x_1, \dots, x_n , 成立

$$p(x_1, \cdots, x_n) = p_1(x_1) \cdots p_n(x_n)$$

C

3.3 随机变量的函数及其分布

3.3.1 Borel 函数与随机变量的函数

定义 3.3.1 (Borel 函数)

称函数 $f: \mathbb{R}^n \to \mathbb{R}$ 为 Borel 函数,如果

$$f^{-1}(\mathscr{B}_1) \subset \mathscr{B}_n$$

命题 3.3.1 (随机变量在 Borel 函数下的像为随机变量)

对于概率空间 (Ω, Σ, P) ,如果 (ξ_1, \dots, ξ_n) 为随机向量,且 $f: \mathbb{R}^n \to \mathbb{R}$ 为 Borel 函数,那么 $f(\xi_1, \dots, \xi_n)$ 为随机变量。

定理 3.3.1 (随机变量的函数的独立性)

如果随机变量 ξ_1, \dots, ξ_n 相互独立,那么对于任意 Borel 函数 $f_k: \mathbb{R} \to \mathbb{R}$,随机变量 $f_1(\xi_1), \dots, f_n(\xi_n)$ 相互独立。

3.3.2 单个随机变量的函数的分布律

目标: 已知随机变量 ξ 的分布函数 F(x) 或密度函数 p(x),求解 $\eta=g(\xi)$ 的分布函数 G(y) 或密度函数 q(y),即

$$G(y) = \int_{q(x) \le y} p(x) dx$$

命题 3.3.2

若 g(x) 严格单调, 其反函数 $g^{-1}(y)$ 存在连续导函数, 则 $\eta = g(\xi)$ 具有密度函数

$$q(y) = p(q^{-1}(y)) \mid (q^{-1}(y))' \mid$$

命题 3.3.3 (倍数分布)

如果随机变量 ξ 的分布函数为 F(x), 密度函数为 p(x), 那么 $\eta = c\xi$ 的分布函数为

$$G(y) = \begin{cases} F(y/c), & c > 0\\ 1 - F(y/c), & c < 0 \end{cases}$$

密度函数为

$$q(y) = \frac{p(y/c)}{|c|}$$

命题 3.3.4 (平方分布)

如果随机变量 ξ 的分布函数为 F(x), 密度函数为 p(x), 则 $\eta = \xi^2$ 的分布函数为

$$G(y) = \begin{cases} F(\sqrt{y}) - F(-\sqrt{y}), & y > 0\\ 0, & y \le 0 \end{cases}$$

密度函数为

$$q(y) = \begin{cases} \frac{p(\sqrt{y}) + p(-\sqrt{y})}{2\sqrt{y}}, & y > 0\\ 0, & y < 0 \end{cases}$$

3.3.3 多个随机变量的函数的分布律

目标: 已知随机向量 (ξ_1,\dots,ξ_n) 的分布函数 $F(x_1,\dots,x_n)$ 或密度函数 $p(x_1,\dots,x_n)$,求解 $\eta=g(\xi_1,\dots,\xi_n)$ 的分布函数 G(y) 或密度函数 q(y),即

$$G(y) = \int_{g(x_1, \dots, x_n) < y} p(x_1, \dots, x_n) dx_1 \dots dx_n$$

命题 3.3.5 (和的分布)

如果随机向量 (ξ_1, ξ_2) 的密度函数为 $p(x_1, x_2)$, 那么 $\eta = \xi_1 + \xi_2$ 的分布函数为

$$G(y) = \iint\limits_{x_1 + x_2 < y} p(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

若 (ξ_1, ξ_2) 相互独立,则

$$G(y) = \int_{-\infty}^{y} dv \int_{-\infty}^{+\infty} p(u, v - u) du = \int_{-\infty}^{y} dv \int_{-\infty}^{+\infty} p(v - u, u) du$$

此时密度函数为

$$q(y) = \int_{-\infty}^{+\infty} p(x, y - x) dx = \int_{-\infty}^{+\infty} p(y - x, x) dx$$

命题 3.3.6 (商的分布)

如果随机向量 (ξ_1, ξ_2) 的密度函数为 $p(x_1, x_2)$, 那么 $\eta = \xi_1/\xi_2$ 的分布函数为

$$G(y) = \iint_{x_1/x_2 < y} p(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

若 (ξ_1, ξ_2) 相互独立,则

$$G(y) = \int_0^{+\infty} dz \int_{-\infty}^{zy} p(x, z) dx + \int_{-\infty}^0 dz \int_{zy}^{+\infty} p(x, z) dx$$

此时密度函数为

$$q(y) = \int_{-\infty}^{+\infty} |z| p(zy, z) dz$$

命题 3.3.7 (顺序统计量的分布)

假设随机变量 ξ_1, \dots, ξ_n 相互独立,且具有相同的分布函数 F(x) 和密度函数 p(x)。

1. 极小值的分布函数与密度函数:

$$G(x) = 1 - (1 - F(x))^n, q(x) = np(x)(1 - F(x))^{n-1}$$

2. 极大值的分布函数与密度函数:

$$G(x) = G(x) = (F(x))^n, q(x) = np(x)(F(x))^{n-1}$$

3. 极小值与极大值的联合分布函数与联合密度函数:

$$G(x,y) = \begin{cases} (F(y))^n, & x \ge y \\ (F(y))^n - (F(y) - F(x))^n, & x < y \end{cases}$$

$$q(x,y) = \begin{cases} 0, & x \ge y \\ n(n-1)q(x)q(y)(F(y) - F(x))^{n-2}, & x < y \end{cases}$$

定理 3.3.2 (随机向量的变换的分布)

如果随机向量 (ξ_1, \dots, ξ_n) 的密度函数为 $p(x_1, \dots, x_n)$, 随机变量 η_1, \dots, η_n

$$\eta_1 = g_1(\xi_1, \dots, \xi_n) \qquad \dots \qquad \eta_n = g_n(\xi_1, \dots, \xi_n)$$

的密度函数为 $q(y_1,\cdots,y_n)$,且对于 $y_k=g_k(x_1,\cdots,x_n)$,存在且存在唯一的反函数 $x_k=x_k(y_1,\cdots,y_n)$,那么

$$q(y_1, \dots, y_n) = p(x_1(y_1, \dots, y_n), \dots, x_n(y_1, \dots, y_n))|J|$$

其中 J 为 Jacobi 行列式

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

 \Diamond

第四章 数字特征与特征函数

4.1 数学期望

定义 4.1.1 (离散型随机变量的数学期望)

对于离散型随机变量 ξ , 令 $P\{\xi = x_k\} = p_k$, 称

$$E(\xi) = \sum_{k=1}^{\infty} x_k p_k$$

为ξ的数学期望,如果该级数绝对收敛。

定义 4.1.2 (连续型随机变量的数学期望)

对于离散型随机变量 ξ , 其密度函数为 p(x), 称

$$E(\xi) = \int_{-\infty}^{\infty} x p(x) \mathrm{d}x$$

为 ξ 的数学期望,如果该积分绝对可积。

定义 4.1.3 (数学期望)

1. 如果随机变量 ξ 的分布函数为 F(x), 且积分

$$\int_{-\infty}^{\infty} x \mathrm{d}F(x)$$

绝对可积, 那么定义

$$E(\xi) = \int_{-\infty}^{\infty} x \mathrm{d}F(x)$$

为ξ的数学期望。

2. 随机向量 (ξ_1, \dots, ξ_n) 的数学期望为 $(E(\xi_1), \dots, E(\xi_n))$, 其中

$$E(\xi_k) = \int_{-\infty}^{\infty} x_k dF_k(x_k), \qquad k = 1, \dots, n$$

命题 4.1.1 (随机变量的函数的数学期望)

1. 如果随机变量 ξ 的分布函数为 F(x), 且 $g: \mathbb{R} \to \mathbb{R}$ 为 Borel 函数,那么随机变量 $\eta = g(\xi)$ 的数学期望为

$$E(\eta) = \int_{-\infty}^{\infty} g(x) dF(x)$$

2. 如果随机向量 (ξ_1,\dots,ξ_n) 的分布函数为 $F(x_1,\dots,x_n)$, 且 $g:\mathbb{R}^n\to\mathbb{R}$ 为 Borel 函数,那么随机变量 $\eta=g(\xi_1,\dots,\xi_n)$ 的数学期望为

$$E(\eta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n) dF(x_1, \cdots, x_n)$$

命题 4.1.2 (数学期望的性质)

- 1. 若 $a \le \xi \le b$, 则 $a \le E(\xi) \le b$ 。
- 2. 线性性:

$$E(\xi + \eta) = E(\xi) + E(\eta), \qquad E(c\xi) = cE(\xi)$$

定理 4.1.1 (Cauchy-Schwarz 不等式)

对随机变量 ξ 和 η , 成立

$$(E(\xi\eta))^2 \le E(\xi^2)E(\eta^2)$$

当且仅当存在常数 a,b, 使得成立

$$P\{a\xi = b\eta\} = 1$$

时等号成立。

 \bigcirc

4.2 方差,相关系数,矩

4.2.1 方差

定义 4.2.1 (方差)

对于随机变量 ξ , 如果

$$D(\xi) = E((\xi - E(\xi))^2) = E(\xi^2) - (E(\xi))^2$$

存在, 那么称 $D(\xi)$ 为 ξ 的方差。

•

定义 4.2.2 (标准差)

称随机变量的方差的方根为标准差, 即 $\sqrt{D(\xi)}$ 。

4

定义 4.2.3 (随机变量的标准化)

对于随机变量 ξ , 若其数学期望 $E(\xi)$ 及方差 $D(\xi)$ 均存在, 且 $D(\xi) > 0$, 则可标准化为

$$\xi^* = \frac{\xi - E(\xi)}{D(\xi)}, \qquad E(\xi^*) = 0, \qquad D(\xi^*) = 1$$

命题 4.2.1 (方差的性质)

1. 非线性性质:

$$D(c\xi) = c^2 D(\xi),$$
 $D(\xi + \eta) = D(\xi) + D(\eta) + 2(E(\xi\eta) - E(\xi)E(\eta))$

2. $D(\xi) \le E(\xi - c)^2$, 当且仅当 $E(\xi) = c$ 时等号成立。

定理 4.2.1 (Chebyshev 不等式)

如果随机变量 ξ 存在数学期望 $E(\xi)$ 与方差 $D(\xi)$, 那么对于任意 $\varepsilon>0$, 成立

$$P\{|\xi - E(\xi)| \ge \varepsilon\} \le \frac{D(\xi)}{\varepsilon^2}$$

 \sim

4.2.2 协方差

定义 4.2.4 (协方差)

对于随机变量 ξ 和 η , 如果

$$cov(\xi, \eta) = E((\xi - E(\xi))(\eta - E(\eta))) = E(\xi\eta) - E(\xi)E(\eta)$$

存在, 那么称 $cov(\xi, \eta)$ 为 ξ 与 η 的协方差。

命题 4.2.2 (协方差的性质)

- 1. 与方差的关系: $cov(\xi,\xi) = D(\xi)$
- 2. 对称性: $cov(\xi, \eta) = cov(\eta, \xi)$
- 3. 线性性:

$$cov(\xi, c\eta) = ccov(\xi, \eta), \qquad cov(\xi, \eta + \zeta) = cov(\xi, \eta) + cov(\xi, \zeta)$$

定义 4.2.5 (协方差矩阵)

定义随机向量 $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$ 的协方差矩阵为

$$D(\boldsymbol{\xi}) = \begin{pmatrix} \cos(\xi_1, \xi_1) & \cdots & \cos(\xi_1, \xi_n) \\ \vdots & \ddots & \vdots \\ \cos(\xi_n, \xi_1) & \cdots & \cos(\xi_n, \xi_n) \end{pmatrix}$$

4.2.3 相关系数与相关性

定义 4.2.6 (相关系数)

1. 对于随机变量 ξ 与 η ,定义其相关系数为

$$\rho = \begin{cases} \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}, &, & \xi \leq \eta \text{ js} \\ 0, & \xi \leq \eta \text{ js} \end{cases}$$

2. 对于事件 A 和 B, 定义其相关系数为

$$\rho = \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)P(A^c)P(B)P(B^c)}}$$

命题 4.2.3 (相关系数的性质)

- 1. 对称性: $\rho(\xi, \eta) = \rho(\eta, \xi)$
- 2. $\rho(\xi, c\eta) = \operatorname{sgn}(c)\rho(\xi, \eta)$

命题 4.2.4 (相关系数的相关性)

令随机变量 ξ 和 η 的相关系数为 ρ 。

- 1. $\rho > 0$: 正相关
- 2. ρ < 0: 负相关
- 3. $\rho = 1$: 完全正相关
- 4. $\rho = -1$: 完全负相关
- 5. ρ = 0: 不相关

命题 4.2.5

对于随机变量 ξ 和 η 的相关系数 ρ ,成立

$$|\rho| \leq 1$$

并且 $\rho=1$ 当且仅当

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

 $\rho = -1$ 当且仅当

$$P\left\{\frac{\xi - E(\xi)}{\sqrt{D(\xi)}} + \frac{\eta - E(\eta)}{\sqrt{D(\eta)}} = 0\right\} = 1$$

定义 4.2.7 (相关性)

称随机变量 ξ 和 η 不相关,如果成立如下命题之一。

- 1. $\rho = 0$
- 2. $cov(\xi, \eta) = 0$
- 3. $E(\xi \eta) = E(\xi)E(\eta)$
- 4. $D(\xi + \eta) = D(\xi) + D(\eta)$

命题 4.2.6 (独立性与相关性)

- 1. 如果随机变量 ξ 和 η 独立, 那么 ξ 和 η 不相关。
- 2. 对于二元正态分布,不相关性与独立性是等价的。
- 3. 对于二值随机变量,不相关性与独立性是等价的。

4.2.4 矩

定义 4.2.8 (原点矩)

定义随机变量 ξ 的n阶原点矩为

$$m_n = E(\xi^n)$$

定义 4.2.9 (中心矩)

定义随机变量 ξ 的 n 阶中心矩为

$$c_n = E((\xi - E(\xi))^n)$$

命题 4.2.7 (原点矩与中心距的关系)

$$c_n = \sum_{k=0}^n \binom{n}{k} (-m_1)^{n-k} m_k, \qquad m_n = \sum_{k=0}^n \binom{n}{k} c_{n-k} m_1^k$$

4.2.5 分位数

定义 **4.2.10** (p 分位数)

对于 $0 , 称<math>x_p$ 为分布函数F(x) 的p 分位数, 如果

$$F(x_p) \le p \le F(x_p^+)$$

定义 4.2.11 (中位数)

称 $x_{0.5}$ 为中位数。

4.2.6 条件数学期望与最小二乘回归

定义 4.2.12 (条件数学期望)

对于随机向量 (ξ, η) , 定义 η 在 $\xi = x$ 的条件下的条件数学期望为

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

命题 4.2.8 (重期望公式)

$$E(\eta) = E(E(\eta \mid \xi))$$

定义 4.2.13 (最小二乘法)

对于随机变量 ξ 和 η ,优化问题

$$\inf_{h} E(\eta - h(\xi))$$

的解为

$$h(x) = E\{\eta \mid \xi = x\}$$

定义 4.2.14 (最小二乘回归)

称

$$y = E\{\eta \mid \xi = x\}$$

为η关于ξ的最小二乘回归。

定义 4.2.15 (线性回归)

如果 ξ 和 η 的数学期望分别为 μ_1 和 μ_2 ,标准差分别为 σ_1 和 σ_2 及相关系数为 ρ ,那么 η 关于 ξ 的线性回归为

$$L(x) = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1)$$

其均方误差为

$$E(\eta - L(\xi)^2) = \sigma_2^2 (1 - \rho^2)$$

4.3 母函数

定义 4.3.1 (母函数)

对于离散型随机变量 ξ , 概率分布为 $P\{\xi=n\}=p_n$, 定义 ξ 的母函数为

$$P(s) = \sum_{n=0}^{\infty} p_n s^n$$

命题 4.3.1 (母函数的性质)

- 1. 唯一性: 母函数与概率分布函数一一对应。
- 2. 数学期望: $E(\xi) = P'(1)$
- 3. 方差: $D(\xi) = P''(1) + P'(1) (P'(1))^2$

命题 4.3.2 (独立随机变量之和的母函数)

若随机变量 ξ 和 η 相互独立, 其相应的母函数分别为 A(s) 和 B(s), 则随机变量 $\xi + \eta$ 的母函数 C(s) 为

$$C(s) = A(s)B(s)$$

命题 4.3.3 (随机个随机变量之和的母函数)

设 ξ_1,ξ_2,\cdots 是相互独立的具有相同概率分布 $P\{\xi_i=j\}=f_j$ 的随机变量,其母函数为

$$F(s) = \sum_{n=0}^{\infty} f_n s^n$$

随机变量 ν 取整数值,且 $P\{\nu=n\}=g_n$,其母函数为

$$G(s) = \sum_{n=0}^{\infty} g_n s^n$$

若 $\{\xi_n\}$ 与 ν 独立,则随机变量 $\eta = \xi_1 + \cdots + \xi_{\nu}$ 的母函数为

4.4 特征函数

4.4.1 特征函数

定义 4.4.1 (特征函数)

对于分布函数为 F(x) 的随机变量 ξ , 定义其特征函数为

$$f(t) = E(e^{it\xi}) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

定义 4.4.2 (连续型随机变量的特征函数)

对于密度函数为 p(x) 的连续型随机变量 ξ , 定义其特征函数为

$$f(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx$$

命题 4.4.1 (特征函数的性质)

1.

$$f(0) = 1,$$
 $|f(t)| \le f(0),$ $f(-t) = \overline{f(t)}$

- 2. 一致连续性: 特征函数 f(t) 在 \mathbb{R} 上一致连续。
- 3. 非负定性:对于任意 $n \in \mathbb{N}^*$ 与实数 t_1, \dots, t_n 及复数 $\lambda_1, \dots, \lambda_n$, 成立

$$\sum_{i,j=1}^{n} f(t_i - t_j) \lambda_i \overline{\lambda_j} \ge 0$$

4. 随机变量和的特征函数: 若随机变量 ξ 和 η 相互独立,且对应的特征函数分别为 $f_{\xi}(t)$ 和 $f_{\eta}(t)$,则 其和 $\xi + \eta$ 的特征函数 $f_{\xi+\eta}(t)$ 满足

$$f_{\xi+\eta}(t) = f_{\xi}(t) f_{\eta}(t)$$

5. 可微性: 若随机变量 ξ 存在 n 阶矩,则其特征函数可微分 n 次,且当 $k \le n$ 时,成立

$$f^{(k)}(0) = i^k E(\xi^k)$$

6. 位移性:

$$f_{a\xi+b}(t) = e^{ibt} f_{\xi}(at)$$

4.4.2 特征函数的唯一性

定理 4.4.1 (逆转公式)

若分布函数 F(x) 的特征函数为 f(t), 且 x_1, x_2 是 F(x) 的连续点,则成立逆转公式

$$F(x_2) - F(x_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itx_1} - e^{-itx_2}}{it} f(t) dt$$

 \sim

定理 4.4.2 (唯一性定理)

分布函数 F(x) 由特征函数 f(t) 唯一确定,且

$$F(x) = \frac{1}{2\pi} \lim_{y \to \infty} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-ity} - \mathrm{e}^{-itx}}{it} f(t) \mathrm{d}t$$

特别的, 若 f(t) 绝对可积, 则相应的分布函数 F(x) 存在并连续, 且

$$F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t) dt$$

C

第五章 极限定理

5.1 Bernoulli 试验场合的极限定理

5.1.1 大数定律与中心极限定理

定义 5.1.1 (大数定律)

称随机变量序列 $\{\xi_n\}_{n=1}^\infty$ 服从大数定律,如果存在数列 $\{a_n\}_{n=1}^\infty$,使得对于任意 $\varepsilon>0$,成立

$$\lim_{n \to \infty} P\{|\eta_n - a_n| < \varepsilon\} = 1$$

其中

$$\eta_n = \frac{\xi_1 + \dots + \xi_n}{n}, \quad n \in \mathbb{N}^*$$

定义 5.1.2 (强大数定律)

称独立随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$ 服从强大数定律,如果

$$P\left\{\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\xi_k - E(\xi_k)) = 0\right\} = 1$$

定义 5.1.3 (中心极限定理)

称存在数学期望与方差的相互独立的随机变量序列 $\{\xi_n\}_{n=1}^\infty$ 服从中心极限定理,如果成立

$$\lim_{n \to \infty} P\{\zeta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \qquad x \in \mathbb{R}$$

其中

$$\zeta_n = \frac{\sum_{k=1}^n \xi_k - \sum_{k=1}^n E(\xi_k)}{\sqrt{\sum_{k=1}^n D(\xi_k)}}, \quad n \in \mathbb{N}^*$$

5.1.2 大数定律

定理 5.1.1 (Chebyshev 大数定律)

如果互不相关的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$ 的方差一致有界, 那么对于任意 $\varepsilon > 0$, 成立

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{k=1}^{n} \xi_k - \frac{1}{n} \sum_{k=1}^{n} E(\xi_k) \right| < \varepsilon \right\} = 1$$

定理 5.1.2 (Markov 大数定理)

对于随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 如果

$$\lim_{n \to \infty} D\left(\frac{\xi_1 + \dots + \xi_n}{n}\right) = 0$$

那么对于任意 $\varepsilon > 0$, 成立

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{k=1}^{n} \xi_k - \frac{1}{n} \sum_{k=1}^{n} E(\xi_k) \right| < \varepsilon \right\} = 1$$

定理 5.1.3 (Bernoulli 大数定律)

如果 μ_n 为 n 次 Bernoulli 试验中某事件出现的次数,p 为该事件在每次试验中出现的概率,那么对于任意 $\varepsilon > 0$,成立

$$\lim_{n \to \infty} P\left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1$$

定理 5.1.4 (Poisson 大数定律)

如果在一个独立实验序列中,某事件在第 k 次试验中出现的概率为 p_k ,以 μ_n 记前 n 次试验中该事件出现的次数,那么对于任意的 $\varepsilon>0$,恒成立

$$\lim_{n \to \infty} P\left\{ \left| \frac{\mu_n}{n} - \frac{p_1 + \dots + p_n}{n} \right| < \varepsilon \right\} = 1$$

5.1.3 De Moivre—Laplace 极限定理

定理 5.1.5 (De Moivre—Laplace 极限定理)

1. 局部极限定理: 如果 μ_n 是 n 次 Bernoulli 试验中某事件出现的次数,令 $x_k=\frac{k-np}{\sqrt{np(1-p)}}$,其中 0< p<1,且满足 $a\leq x_k\leq b$,那么

$$P\{\mu_n = k\} \sim \frac{1}{\sqrt{np(1-p)}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_k^2}{2}} \qquad (n \to \infty)$$

2. 积分极限定理: 如果 μ_n 是 n 次 Bernoulli 试验中某事件出现的次数, 那么对于 0 , 成立

$$\lim_{n \to \infty} P\left\{ a \le \frac{\mu_n - np}{\sqrt{np(1-p)}} \le b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

推论 5.1.1

1. 用频率估计概率:

$$P\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} \approx \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\varepsilon\sqrt{\frac{n}{p(1-p)}}} e^{-\frac{x^2}{2}} dx - 1$$

2. 二项式分布计算:

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{1}{\sqrt{np(1-p)}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{k-np}{\sqrt{np(1-p)}})^2}$$

3. 二项分布计算:

$$P\{a \le \mu_n \le b\} \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{a-np}{\sqrt{np(1-p)}}}^{\frac{b-np}{\sqrt{np(1-p)}}} e^{-\frac{x^2}{2}} dx$$

 \Diamond

5.2 收敛性

5.2.1 分布函数弱收敛

定义 5.2.1 (弱收敛)

称分布函数序列 $\{F_n(x)\}$ 弱收敛于单调递增函数 F(x),并记作 $F_n(x) \xrightarrow{W} F(x)$,如果对于任意 F(x) 的连续点 x,成立

$$\lim_{n \to \infty} F_n(x) = F(x)$$

在 F(x) 的每一连续点上都成立,则称 $F_n(x)$,。

定理 5.2.1 (Helly 第一定理)

一致有界的单调递增函数序列 $\{F_n(x)\}$ 存在子序列 $\{F_{n_k}(x)\}$ 弱收敛于有界单调递增函数 F(x)。

定理 5.2.2 (Helly 第二定理)

如果 f(x) 是 [a,b] 上的连续函数,且一致有界单调递增函数序列 $\{F_n(x)\}$ 在 [a,b] 上弱收敛于函数 F(x),同时 a 和 b 是 F(x) 的连续点,那么

$$\lim_{n \to \infty} \int_{a}^{b} f(x) dF_{n}(x) = \int_{a}^{b} f(x) dF(x)$$

定理 5.2.3 (Helly 第二定律的推广)

如果 f(x) 是 \mathbb{R} 上的连续有界函数,且一致有界单调递增函数序列 $\{F_n(x)\}$ 在 [a,b] 上弱收敛于函数 F(x),同时

$$\lim_{n \to \infty} F_n(-\infty) = F(-\infty), \qquad \lim_{n \to \infty} F_n(+\infty) = F(+\infty)$$

那么

$$\lim_{n \to \infty} \int_{-\infty}^{+\infty} f(x) dF_n(x) = \int_{-\infty}^{+\infty} f(x) dF(x)$$

5.2.2 Lévy—Cramer 定理

定理 **5.2.4** (Lévy—Cramer 正极限定理)

如果分布函数序列 $\{F_n(x)\}$ 弱收敛于分布函数 F(x),那么 $\{F_n(x)\}$ 的特征函数序列 $\{f_n(t)\}$ 内闭一致收敛于 F(x) 的特征函数 f(t)。

定理 5.2.5 (Lévy—Cramer 逆极限定理)

如果特征函数序列 $\{f_n(t)\}$ 收敛于特征函数 f(t),且 f(t) 在 t=0 处连续,那么 $\{f_n(t)\}$ 的分布函数序列 $\{F_n(x)\}$ 弱收敛于 f(t) 的分布函数 F(x)。

5.2.3 随机变量的收敛性

定义 5.2.2 (依分布收敛)

对于随机变量序列 $\{\xi_n\}$ 与随机变量 ξ ,称 $\{\xi_n\}$ 依分布收敛于 ξ ,并记作 $\xi_n \xrightarrow{L} \xi$,如果 $F_n(x) \xrightarrow{W} F(x)$,其中 $\{\xi_n\}$ 与 ξ 的分布函数分别为 $\{F_n(x)\}$ 与 F(x)。

定义 5.2.3 (依概率收敛)

对于随机变量序列 $\{\xi_n\}$ 与随机变量 ξ ,称 $\{\xi_n\}$ 依概率收敛于 ξ ,并记作 $\xi_n \xrightarrow{P} \xi$,如果对于任意 $\varepsilon > 0$,成立

$$\lim_{n \to \infty} P\{|\xi_n - \xi| \ge \varepsilon\} = 0$$

定义 5.2.4 (r 阶收敛)

对于存在 r 阶矩的随机变量序列 $\{\xi_n\}$ 与随机变量 ξ ,称 $\{\xi_n\}r$ 阶收敛于 ξ ,并记作 $\xi_n \xrightarrow{r} \xi$,如果 $\lim_{n \to \infty} E(|\xi_n - \xi|^r) = 0$

定义 5.2.5 (以概率 1 收敛)

对于存在 r 阶矩的随机变量序列 $\{\xi_n\}$ 与随机变量 ξ ,称 $\{\xi_n\}$ 以概率 1 收敛于 ξ ,并记作 $\xi_n \xrightarrow{\text{a.s.}} \xi$,如果 $P\left\{\lim_{n\to\infty}\xi_n=\xi\right\}=1$

命题 5.2.1

$$\begin{array}{c} \xi_n \stackrel{r}{\longrightarrow} \xi \implies \xi_n \stackrel{\mathrm{P}}{\longrightarrow} \xi \implies \xi_n \stackrel{\mathrm{L}}{\longrightarrow} \xi \\ \xi_n \stackrel{\mathrm{a.s.}}{\longrightarrow} \xi \implies \xi_n \stackrel{\mathrm{P}}{\longrightarrow} \xi \\ \\ \text{对于常数} C, \xi_n \stackrel{\mathrm{P}}{\longrightarrow} C \iff \xi_n \stackrel{\mathrm{L}}{\longrightarrow} C \end{array}$$

5.2.4 Bochner—Khinchin 定理

定理 5.2.6 (Bochner—Khinchin 定理)

函数 f(t) 是特征函数的充分必要条件是: f(t) 非负定, 连续, 且 f(0) = 1。

定理 5.2.7 (Herglotz 定理)

复数列 $\{C_n\}_{n=1}^{\infty}$ 可以表示为

$$C_n = \int_{-\pi}^{\pi} e^{inx} dG(x)$$

的充分必要条件是其为非负定的,即对于任意 $n\in\mathbb{N}^*$ 及 $\{\lambda_n\}_{n=1}^\infty\subset\mathbb{C}$ 均有

$$\sum_{i,j=1}^{n} C_{i-j} \lambda_i \overline{\lambda_j} \ge 0$$

其中 G(x) 是 $[-\pi,\pi]$ 上有界、单调非减、左连续函数。

5.3 独立同分布场合的极限定理

5.3.1 Khinchin 大数定律

定理 5.3.1 (Khinchin 大数定律)

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 如果 ξ_n 服从相同的分布, 且

$$E(\xi_n) = \mu, \qquad n \in \mathbb{N}^*$$

那么对于任意 $\varepsilon > 0$, 成立

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{k=1}^{n} \xi_k - \mu \right| < \varepsilon \right\} = 1$$

5.3.2 中心极限定理

定理 5.3.2 (Lindeberg—Lévy 中心极限定理)

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 如果 ξ_n 服从相同的分布,且

$$E(\xi_n) = \mu, \qquad D(\xi_n) = \sigma^2, \qquad n \in \mathbb{N}^*$$

那么标准化随机变量

$$\zeta_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (\xi_k - \mu)$$

成立

$$\lim_{n \to \infty} P\{\zeta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

推论 5.3.1

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 如果 ξ_n 服从相同的分布, 且

$$E(\xi_n) = \mu, \qquad D(\xi_n) = \sigma^2, \qquad n \in \mathbb{N}^*$$

那么

$$P\left\{a \le \sum_{k=1}^{n} \xi_k \le b\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{a-n\mu}{\sigma\sqrt{n}}}^{\frac{b-n\mu}{\sigma\sqrt{n}}} e^{-\frac{t^2}{2}} dt$$

5.4 强大数定律

5.4.1 Borel 强大数定律

定理 5.4.1 (Borel—Cantelli 引理)

1. 对于随机事件序列 $\{A_n\}_{n=1}^{\infty}$, 如果

$$\sum_{n=1}^{\infty} P(A_n) < \infty, \qquad n \in \mathbb{N}^*$$

那么

$$P\left\{\limsup_{n\to\infty} A_n\right\} = 0, P\left\{\liminf_{n\to\infty} A^c\right\} = 1$$

2. 对于相互独立的随机事件序列 $\{A_n\}_{n=1}^{\infty}$, 成立

$$\sum_{n=1}^{\infty}P(A_n)=\infty, \qquad n\in\mathbb{N}^*\iff P\left\{\limsup_{n\to\infty}A_n\right\}=0 \not \lesssim P\left\{\liminf_{n\to\infty}A_n^c\right\}=1$$

定理 5.4.2 (Borel 强大数定律)

如果 μ_n 为某事件在 n 次独立试验中出现的次数,且在每次试验中该事件出现的概率均为 p,那么

$$\lim_{n \to \infty} P\left\{\frac{\mu_n}{n} \to p\right\} = 1$$

 $^{\circ}$

5.4.2 Kolmogorov 强大数定律

定理 5.4.3 (Kolmogorov 不等式)

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^\infty$, 如果对于任意 $n\in\mathbb{N}^*$, $D(\xi_n)$ 存在且有限,那么对于任意 $\varepsilon>0$, 成立

$$P\left\{\max_{1\leq k\leq n}\left|\sum_{i=1}^{k}\left(\xi_{i}-E(\xi_{i})\right)\right|\geq\varepsilon\right\}\leq\frac{1}{\varepsilon^{2}}\sum_{k=1}^{n}D(\xi_{k})$$

m

定理 5.4.4 (Hājek—Rényi 不等式)

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 如果对于任意 $n \in \mathbb{N}^*$, $D(\xi_n) = \sigma_n^2 < \infty$, 且 $\{C_n\}_{n=1}^{\infty} \subset \mathbb{R}^+$ 为单调递减序列,那么对于任意 $m < n \in \mathbb{N}^*$ 与 $\varepsilon > 0$,成立

$$P\left\{\max_{m\leq k\leq n}\left|\sum_{i=1}^{k}\left(\xi_{i}-E(\xi)_{i}\right)\right|\geq\varepsilon\right\}\leq\frac{1}{\varepsilon^{2}}\left(C_{m}^{2}\sum_{i=1}^{m}\sigma_{i}^{2}+\sum_{i=m+1}^{n}C_{i}^{2}\sigma_{i}^{2}\right)$$

定理 5.4.5 (Kolmogorov 强大数定律)

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 如果

$$\sum_{n=1}^{\infty} \frac{D(\xi_n)}{n^2} < \infty$$

那么

$$P\left\{\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\xi_k - E(\xi_k)) = 0\right\} = 1$$

 \sim

定理 5.4.6 (独立同分布场合的强大数定律)

对于相互独立且服从相同分布的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 成立

$$\frac{1}{n} \sum_{k=1}^{n} \xi_k \xrightarrow{\text{a.s.}} \mu \iff E(\xi_n) = \mu, n \in \mathbb{N}^*$$

 \sim

5.5 中心极限定理

目标:对于相互独立且服从相同分布的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$,作标准化

$$\zeta_n = \frac{\sum_{k=1}^n \xi_k - \sum_{k=1}^n E(\xi_k)}{\sqrt{\sum_{k=1}^n D(\xi_k)}}, \quad n \in \mathbb{N}^*$$

本节寻找 ζ_n 的分布函数为正态分布函数的充分必要条件。

定义 5.5.1 (Lindeberg 条件)

对于随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 令

$$\mu_n = E(\xi_n), \quad \sigma_n^2 = D(\xi_n), \quad \Sigma_n^2 = \sum_{k=1}^n \sigma_k^2, \quad n \in \mathbb{N}^*$$

对于任意 $\tau > 0$, 成立

$$\lim_{n \to \infty} \frac{1}{\Sigma_n^2} \sum_{k=1}^n \int_{|x-\mu_j| > \tau \Sigma_n} (x - \mu_k)^2 dF_k(x) = 0$$

其中 $F_n(x)$ 为 ξ_n 的分布函数。

定义 5.5.2 (Felelr 条件)

对于随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 令

$$\mu_n = E(\xi_n), \qquad \sigma_n^2 = D(\xi_n), \qquad \Sigma_n^2 = \sum_{k=1}^n \sigma_k^2, \qquad n \in \mathbb{N}^*$$

成立

$$\lim_{n\to\infty} \max_{1\leq k\leq n} \frac{\sigma_k}{\Sigma_n} = 0 \iff \lim_{n\to\infty} \Sigma_n = \infty \mathbb{H} \lim_{n\to\infty} \frac{\sigma_n}{\Sigma_n} = 0$$

定理 5.5.1 (Lindeberg—Felelr 定理)

对于相互独立且服从相同分布的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 作标准化

$$\zeta_n = \frac{\sum_{k=1}^n \xi_k - \sum_{k=1}^n E(\xi_k)}{\sqrt{\sum_{k=1}^n D(\xi_k)}}, \quad n \in \mathbb{N}^*$$

那么成立

$$\lim_{n \to \infty} P\{\zeta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

与 Felelr 条件的充分必要条件是成立 Lindeberg 条件。

推论 5.5.1

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^\infty$,作标准化

$$\zeta_n = \frac{\sum_{k=1}^n \xi_k - \sum_{k=1}^n E(\xi_k)}{\sqrt{\sum_{k=1}^n D(\xi_k)}}, \quad n \in \mathbb{N}^*$$

令

$$\mu_n = E(\xi_n), \qquad \sigma_n^2 = D(\xi_n), \qquad \Sigma_n^2 = \sum_{k=1}^n \sigma_k^2, \qquad n \in \mathbb{N}^*$$

如果存在常数 K_n , 使得成立

$$\max_{1 \le k \le n} |\xi_k| \le K_n, \qquad n \in \mathbb{N}^*$$

且

$$\lim_{n \to \infty} \frac{K_n}{\Sigma_n} = 0$$

那么

$$\lim_{n \to \infty} P\{\zeta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

推论 5.5.2 (Lyapunov 定理)

对于相互独立的随机变量序列 $\{\xi_n\}_{n=1}^{\infty}$, 作标准化

$$\zeta_n = \frac{\sum_{k=1}^n \xi_k - \sum_{k=1}^n E(\xi_k)}{\sqrt{\sum_{k=1}^n D(\xi_k)}}, \quad n \in \mathbb{N}^*$$

令

$$\mu_n = E(\xi_n), \qquad \sigma_n^2 = D(\xi_n), \qquad \Sigma_n^2 = \sum_{k=1}^n \sigma_k^2, \qquad n \in \mathbb{N}^*$$

如果存在 $\delta > 0$, 使得成立

$$\lim_{n \to \infty} \frac{1}{\sum_{k=1}^{n}} \sum_{k=1}^{n} E(|\xi_k - \mu_k|^{2+\delta}) = 0$$

那么

$$\lim_{n \to \infty} P\{\zeta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

 \sim

附录 A 收敛性, 大数定律与中心极限定理

A.1 收敛性

A.1.1 分布函数

定义 A.1.1 (Lebesgue-Stieltjes 函数)

称函数 $F: \mathbb{R}^n \to \mathbb{R}$ 为 Lebesgue-Stieltjes 函数,如果 F 在 \mathbb{R}^n 上处处上连续,且 F 在任一方体 (a,b] 上具有非负增量。

定义 A.1.2 (分布函数 distributio function)

称 Lebesgue-Stieltjes 函数 $F: \mathbb{R}^n \to \mathbb{R}$ 为分布函数,如果 F 在 \mathbb{R}^n 上单调递增,且对于任意 $1 \le k \le n$,成立

$$F(x_1, \dots, x_{k-1}, -\infty, x_{k+1}, \dots, x_n) = 0, \qquad F(+\infty) = 1$$

A.1.2 随机变量的四大收敛性

定义 A.1.3 (依 L^p 收敛 converge in L^p)

对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$ 与随机变量 X,称 $\{X_n\}_{n=1}^{\infty}$ 依 L^p 收敛于 X,并记作 $X_n \xrightarrow{L^p} X$,如果

$$\lim_{n \to \infty} E(|X_n - X|^p) = 0$$

定义 A.1.4 (几乎必然收敛 converge almost surely)

对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$ 与随机变量 X,称 $\{X_n\}_{n=1}^{\infty}$ 几乎必然收敛于 X,并记作 $X_n \xrightarrow{\text{a.s.}} X$,如果对于任意 $\varepsilon > 0$,成立

$$P\left\{\lim_{n\to\infty}|X_n-X|\right\}=1$$

定义 A.1.5 (依概率收敛 converge in probability)

对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$ 与随机变量 X,称 $\{X_n\}_{n=1}^{\infty}$ 依概率收敛于 X,并记作 $X_n \stackrel{\mathrm{P}}{\longrightarrow} X$,如果对于 任意 $\varepsilon > 0$,成立

$$\lim_{n \to \infty} P\{|X_n - X| \ge \varepsilon\} = 0$$

定义 A.1.6 (弱收敛 converge weakly)

对于有界 Lebesgue-Stieltjes 函数序列 $\{F_n: \mathbb{R} \to \mathbb{R}\}_{n=1}^{\infty}$ 与有界 Lebesgue-Stieltjes 函数 F,称 $\{F_n: \mathbb{R} \to \mathbb{R}\}_{n=1}^{\infty}$ 弱收敛于 F,并记作 $F_n \xrightarrow{w} F$,如果对于任意 F 的连续点 x,成立 $F_n(x) \to F(x)$,且 $F_n(\pm \infty) \to F(\pm \infty)$ 。

定义 A.1.7 (依分布收敛 converge in distribution)

对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$ 与随机变量 X, F_n 为 X_n 的分布函数, F 为 X 的分布函数, 称 $\{X_n\}_{n=1}^{\infty}$ 依 分布收敛于 X, 并记作 $X_n \xrightarrow{d} X$, 如果 $F_n \xrightarrow{w} F$ 。

命题 A.1.1

$$X_n \xrightarrow{L^p} X \implies X_n \xrightarrow{P} X \implies X_n \xrightarrow{d} X, \qquad X_n \xrightarrow{d} X \implies X_n \xrightarrow{P} X$$

A.2 大数定律

A.2.1 大数定理的定义

定义 A.2.1 (大数定律 weak law of large numbers)

1. 对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和,称 $\{X_n\}_{n=1}^{\infty}$ 服从古典意义下的大数定律,如果诸期望 $E(X_n)$ 存在且有限,同时

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

2. 对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和,称 $\{X_n\}_{n=1}^{\infty}$ 服从现代意义下的大数定律,如果存在中心化数列 $\{a_n\}_{n=1}^{\infty}$ 与正则化数列 $\{b_n\}_{n=1}^{\infty}$, 其中 $0 < b_n \to \infty$,使得成立

$$\frac{S_n - a_n}{b_n} \xrightarrow{P} 0$$

A.2.2 经典大数定律

定理 A.2.1 (Khinchin 大数定律)

对于随机变量序列 $\{X_n\}_{n=1}^\infty$, $S_n=\sum_{k=1}^n X_k$ 为其部分和,如果 $\{X_n\}_{n=1}^\infty$ 相互独立且同分布于随机变量 X,同时存在有限期望 E(X),那么

$$\frac{S_n}{n} \xrightarrow{\mathrm{P}} E(X)$$

定理 A.2.2 (Markov 大数定律)

对于随机变量序列 $\{X_n\}_{n=1}^\infty$, $S_n=\sum_{k=1}^n X_k$ 为其部分和,如果诸方差 $D(X_n)$ 存在且有限,且存在实数列 $\{b_n\}_{n=1}^\infty$,使得成立 Markov 条件

$$\frac{D(S_n)}{b_n^2} \to 0$$

那么

$$\frac{S_n - E(S_n)}{b_n} \stackrel{P}{\longrightarrow} 0$$

定理 A.2.3 (Chebyshëv 大数定律)

对于随机变量序列 $\{X_n\}_{n=1}^\infty$, $S_n=\sum_{k=1}^n X_k$ 为其部分和,如果 $\{X_n\}_{n=1}^\infty$ 互不相关且方差一致有界,那么

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

C

定理 A.2.4 (Poisson 大数定律)

对于随机变量序列 $\{X_n\}_{n=1}^\infty$, $S_n=\sum_{k=1}^n X_k$ 为其部分和, 如果 $\{X_n\}_{n=1}^\infty$ 相互独立且

$$P\{X_n = 1\} = p_n \in (0, 1), \qquad P\{X_n = 0\} = 1 - p_n, \qquad n \in \mathbb{N}^*$$

那么

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

定理 A.2.5 (Bernoulli 大数定律)

在每次成功概率为p的 Bernoulli 试验序列中,若以 μ_n 表示前n次试验中成功的次数,那么

$$\frac{\mu_n}{n} \xrightarrow{P} p$$

表 A.1: 经典大数定律

名称	分布	期望	方差	相关性	独立性	条件	结论
Khinchin 大数定律	同分布于 X	E(X) 存在且有限			相互独立		$\xrightarrow{\underline{S_n}} \xrightarrow{\mathrm{P}} E(X)$
Markov 大数定理		$E(X_n)$ 存在且有限	$D(X_n)$ 存在且有限			$\frac{D(S_n)}{b_n^2} \to 0$	$\frac{S_n - E(S_n)}{b_n} \xrightarrow{\mathbf{P}} 0$
Chebyshëv 大数定律		$E(X_n)$ 存在且有限	$D(X_n)$ 存在且有限	互不相关		$D(X_n)$ 一致有界	$\xrightarrow[n]{S_n - E(S_n)} \xrightarrow[n]{\mathbf{P}} 0$
Poisson 大数定律	$P\{X_n = 1\} = p_n, P\{X_n = 0\} = 1 - p_n$				相互独立		$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$
Bernoulli 大数定律	每次成功概率为 p 的 Bernoulli 试验序列					以 μ_n 表示前 n 次试验中成功的次数	$\frac{\mu_n}{n} \xrightarrow{\mathcal{P}} p$

A.3 强大数定律

A.3.1 强大数定律的定义

定义 A.3.1 (强大数定律 strong law of large numbers)

1. 对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和,称 $\{X_n\}_{n=1}^{\infty}$ 服从古典意义下的强大数定律,如果诸期望 $E(X_n)$ 存在且有限,同时

$$\frac{S_n - E(S_n)}{n} \xrightarrow{\text{a.s.}} 0$$

2. 对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和,称 $\{X_n\}_{n=1}^{\infty}$ 服从现代意义下的强大数定律,如果存在中心化数列 $\{a_n\}_{n=1}^{\infty}$ 与正则化数列 $\{b_n\}_{n=1}^{\infty}$,其中 $0 < b_n \to \infty$,使得成立

$$\frac{S_n - a_n}{b_n} \xrightarrow{\text{a.s.}} 0$$



A.3.2 经典强大数定律

定理 A.3.1 (Kolmogorov 强大数定律)

对于相互独立且同分布于 X 的随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和,成立如下命题。

1. 若 E(|X|) 存在且有限, 那么

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} E(X)$$

2. 若 $\frac{S_n}{n} \xrightarrow{\text{a.s.}} \mu$, 那么 E(|X|) 存在且有限, 同时 $E(X) = \mu$ 。

\Diamond

定理 A.3.2 (Marcinkiewicz-Zygmund 强大数定律)

对于相互独立且同分布于 X 的随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n=\sum_{k=1}^n X_k$ 为其部分和,如下命题等价。

1. 存在常数 $a \in \mathbb{R}$ 与 $p \in (0,2)$, 使得成立

$$\frac{S_n - na}{n^{-\frac{1}{p}}} \xrightarrow{\text{a.s.}} 0$$

2. p 阶矩 $E(|X|^p)$ 存在且有限。

此时当0 时,<math>a可取任意常数;当 $1 \le p < 2$ 时,a = E(X)。



定理 A.3.3 (独立但不同分布的强大数定律)

对于相互独立的随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n=\sum_{k=1}^n X_k$ 为其部分和,实数列 $\{b_n\}_{n=1}^{\infty}$ 成立 $b_n>0$ 且单调递增趋于 ∞ ,成立如下命题。

1. 如果存在 0 , 使得诸 <math>p 阶矩 $E(|X|^p)$ 存在, 且

$$\sum_{n=1}^{\infty} \frac{E(|X_n|^p)}{b_n^p} < \infty$$

那么

$$\frac{S_n}{b_n} \xrightarrow{\text{a.s.}} 0$$

2. 如果存在 1 , 使得诸 <math>p 阶矩 $E(|X|^p)$ 存在, 且

$$\sum_{n=1}^{\infty} \frac{E(|X_n|^p)}{b_n^p} < \infty$$

那么

$$\frac{S_n - E(S_n)}{b_n} \xrightarrow{\text{a.s.}} 0$$



A.4 中心极限定理

A.4.1 中心极限定理的定义

定义 A.4.1 (中心极限定理 central limit theorem)

1. 对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^{n} X_k$ 为其部分和, 称 $\{X_n\}_{n=1}^{\infty}$ 服从古典意义下的中心极限定 理,如果诸方差 $D(X_n)$ 存在且有限,同时

$$\frac{S_n - E(S_n)}{\sqrt{D(S_n)}} \stackrel{\mathrm{d}}{\longrightarrow} N(0, 1)$$

2. 对于随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和,称 $\{X_n\}_{n=1}^{\infty}$ 服从现代意义下的中心极限定理,如果存在中心化数列 $\{a_n\}_{n=1}^{\infty}$ 与正则化数列 $\{b_n\}_{n=1}^{\infty}$,其中 $0 < b_n \to \infty$,使得成立

$$\frac{S_n - a_n}{b_n} \xrightarrow{\mathbf{d}} N(0, 1)$$

A.4.2 经典中心极限定理

定理 A.4.1 (Dé Moivre-Laplace 中心极限定理)

对于相互独立且同分布于 X 的随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^{n} X_k$ 为其部分和,如果

$$P{X = 1} = p \in (0, 1), \qquad P{X = 0} = 1 - p$$

那么

$$\frac{S_n - np}{\sqrt{np(1-p)}} \stackrel{\mathrm{d}}{\longrightarrow} N(0,1)$$

定理 A.4.2 (Lindeberg-Lévy 中心极限定理)

对于相互独立且同分布于 X 的随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^n X_k$ 为其部分和, 如果期望 E(X) 存在 且有限, 方差 D(X) 存在且有限, 那么

$$\frac{S_n - nE(X)}{\sqrt{nD(X)}} \xrightarrow{d} N(0,1)$$

定理 A.4.3 (Lindeberg-Feller 中心极限定理)

对于相互独立的随机变量序列 $\{X_n\}_{n=1}^{\infty}$, $S_n = \sum_{k=1}^{n} X_k$ 为其部分和, 诸期望 $E(X_n)$ 存在且有限, 诸方差 $D(X_n)$ 存在且有限,令

$$\mu_n = E(X_n), \qquad \sigma_n^2 = D(X_n), \qquad B_n^2 = \sum_{k=1}^n \sigma_k^2$$

1. Feller 条件: $\lim_{n \to \infty} \max_{1 \le k \le n} \frac{\sigma_k^2}{B_n^2} = 0$ 2. $\frac{S_n - E(S_n)}{B_n} \xrightarrow{d} N(0, 1)$

2.
$$\frac{S_n - E(S_n)}{B} \xrightarrow{d} N(0,1)$$

成立的充分必要条件为 Lindeberg 条件, 即对于任意 $\varepsilon > 0$, 成立

$$\lim_{n \to \infty} \frac{1}{B_n^2} \sum_{k=1}^n E(X_k - \mu_k)^2 I\{|X_k - \mu_k| \ge \varepsilon B_n\} = 0$$

\Diamond

定理 A.4.4 (Lyapunov 中心极限定理)

对于相互独立的随机变量序列 $\{X_n\}_{n=1}^\infty$, $S_n=\sum_{k=1}^n X_k$ 为其部分和,如果 $\{X_n\}_{n=1}^\infty$ 成立 Lyapunov 条件,即存在 $\delta>0$,使得成立

$$\lim_{n \to \infty} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^n E(|X_k - E(X_k)|^{2+\delta}) = 0$$



附录 B 概率模型

概率模型	密度函数 $p(x)$	参数范围	数学期望 <i>E</i> (ξ)	方差 D(ξ)	特征函数 $f(t)$
退化分布 $I_c(x)$	$p(x) = \begin{cases} 1, & x = c \\ 0, & x \neq c \end{cases}$		c	0	e^{ict}
Bernoulli 分布	$p(x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \end{cases}$	0	p	p(1-p)	$pe^{it} + 1 - p$
二项分布 $B(n,p)$	$b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}$	$0 \leq k \leq n; 0$	np	np(1-p)	$(pe^{it} + 1 - p)^n$
Poisson 分布 $P(\lambda)$	$p(k;\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$	$k\in\mathbb{N}; \lambda>0$	λ	λ	$\mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}$
几何分布	$g(k;p) = p(1-p)^{k-1}$	$k \in \mathbb{N}^*, 0$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{p\mathrm{e}^{it}}{1-(1-p)\mathrm{e}^{it}}$
超几何分布	$p_k = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$	$M,n \leq N; 0 \leq k \leq \min\{M,n\}$	$\frac{nM}{N}$	$\frac{nM(N-M)(N-n)}{N^2(N-1)}$	$\sum_{k=0}^{n} \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} e^{ikt}$
Pascal 分布	$p_k = \binom{k-1}{r-1} p^r (1-p)^{k-r}$	$k \geq r, 0$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\big(\frac{(1-p)\mathrm{e}^{it}}{1-(1-p)\mathrm{e}^{it}}\big)^{r}$
负二项分布	$p_k = \binom{-r}{k} p^r (p-1)^k$	$k \in \mathbb{N}, 0 0$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$(\frac{p}{1-(1-p)\mathrm{e}^{it}})^r$
正态分布 $N(\mu, \sigma^2)$	$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$		μ	σ	$\mathrm{e}^{i\mu t - \frac{1}{2}\sigma^2 t^2}$
均匀分布 U[a, b]	$p(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & 其他 \end{cases}$	a < b	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{\mathrm{e}^{ibt} - \mathrm{e}^{iat}}{i(b-a)t}$
指数分布 Exp(λ)	$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0\\ 0, & x < 0 \end{cases}$	$\lambda > 0$	λ^{-1}	λ^{-2}	$(1-\frac{it}{\lambda})^{-1}$
χ ² 分布	$p(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2} - 1} e^{-\frac{x}{2}}, & x \ge 0\\ 0, & x < 0 \end{cases}$	$n\in \mathbb{N}^*$	n	2n	$(1-2it)^{-\frac{n}{2}}$
Γ 分布 $\Gamma(r,\lambda)$	$p(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x \ge 0\\ 0, & x < 0 \end{cases}$	$r, \lambda > 0$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$(1-\frac{it}{\lambda})^{-r}$
Cauchy 分布	$p(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - \mu)^2}$	$\mu \in \mathbb{R}, \lambda > 0$	不存在	不存在	$\mathrm{e}^{i\mu t - \lambda t }$
t 分布	$p(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} (1 + \frac{x^2}{n})^{-\frac{n+1}{2}}$	$n\in\mathbb{N}^*$	0(n > 1)	$\frac{n}{n-2}(n>2)$	
Pareto 分布	$p(x) = \begin{cases} rA^r \frac{1}{x^{r+1}}, & x \ge A \\ 0, & x < A \end{cases}$	r, A > 0	(r > 1 时存在)	(r>2时存在)	
F 分布	$p(x) = \begin{cases} rA^{r} \frac{1}{x^{r+1}}, & x \ge A \\ 0, & x < A \end{cases}$ $p(x) = \begin{cases} \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{m}{2})} m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{x^{\frac{m}{2}-1}}{x^{\frac{m}{2}-1}}, & x \ge 0 \\ 0, & x < 0 \end{cases}$	$m,n\in\mathbb{N}^*$	$\frac{n}{n-2}(n>2)$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}(n>4)$	
β分布	$p(x) = \begin{cases} \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1} (1-x)^{q-1}, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$	p,q>0	$\frac{p}{p+q}$	$\frac{pq}{(p+q)^2(p+q+1)}$	$\frac{\Gamma(p+q)}{\Gamma(p)} \sum_{k=0}^{\infty} \frac{\Gamma(p+k)(it)^k}{\Gamma(p+q+k)\Gamma(k+1)}$
对数正态分布	$p(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{\frac{(\ln x - \alpha)^2}{2\sigma^2}}, & x > 0\\ 0, & x \le 0 \end{cases}$	$\alpha, \sigma > 0$	$e^{\alpha + \frac{\sigma^2}{2}}$	$e^{2\alpha+\sigma^2}(e^{\sigma^2}-1)$	
Weibull 分布		$\lambda, \alpha > 0$	$\Gamma(\frac{1}{\alpha}+1)\lambda^{-\frac{1}{\alpha}}$	$\lambda^{-\frac{2}{\alpha}}(\Gamma(\frac{2}{\alpha}+1)-(\Gamma(\frac{1}{\alpha}+1))^2)$	
Rayleigh 分布	$p(x) = \begin{cases} \alpha \lambda x^{\alpha - 1} e^{-\lambda x^{\alpha}}, & x > 0 \\ 0, & x \le 0 \end{cases}$ $p(x) = \begin{cases} x e^{-\frac{x^{2}}{2}}, & x \ge 0 \\ 0, & x < 0 \end{cases}$		$\sqrt{\frac{\pi}{2}}$	$2-\frac{\pi}{2}$	