

概率论

Fir1247

2023 年 3 月-2023 年 5 月

前言

本文档为 2023 年春季学期张娜老师的《概率论》课程笔记。

为避免歧义，本文档中的 $\subset$ ，相当于 $\subseteq, \subseteq$ ，均表示“包含于”的意思，且只使用第一种形式；本文档中的 $\subsetneq$ ，相当于 $\subsetneq$ ，均表示“真包含于”的意思，且只使用第一种形式。

期中考试范围是 2.2 节及之前。

Fir1247

目录

第一章	事件与概率	1
1.1	基础概念	1
1.2	概率	1
1.3	条件概率	3
1.4	独立性	4
1.5	乘积空间	5
1.6	一些例子	5
第二章	随机变量及分布函数	7
2.1	概念介绍	7
2.1.1	离散型随机变量	8
2.1.2	连续型随机变量	8
2.1.3	随机向量	8
2.2	离散型随机变量	9
2.2.1	常见离散型分布	9
2.2.2	独立性	9
2.2.3	数学期望	9
2.2.4	方差	9
2.2.5	常见随机分布的期望和方差	10
2.2.6	示性函数举例	10
2.2.7	条件分布与条件期望	10
2.2.8	母函数	11
2.2.9	随机游走	13
2.3	连续性随机变量	15
2.3.1	密度函数的独立性	15
2.3.2	数学期望	16
2.3.3	常用连续型分布	17
2.3.4	连续型随机向量	19
2.4	随机变量的函数	22
2.5	多元正态分布	24
第三章	特征函数及其应用	26
3.1	数学期望	26
3.1.1	抽象积分	26
3.2	特征函数	27
3.2.1	常见分布的特征函数	28
3.2.2	反转公式与唯一性定理	29

3.2.3	多元特征函数	29
3.2.4	函数列的收敛性	30
3.3	极限定理简要介绍	31
第四章	极限定理	32
4.1	收敛性	32
4.1.1	收敛性的定义	32
4.1.2	四种收敛性之间的关系	32
4.2	辅助结论	34
4.2.1	不等式	34
4.2.2	运算性质	34
4.2.3	Borel-Cantelli 引理	35
4.3	大数定律	35
4.3.1	弱大数定律	35
4.3.2	强大数定律	36
4.4	中心极限定理	36

第一章 事件与概率

1.1 基础概念

定义 1.1.1 随机实验：可重复、所有可能结果已知、每次实验结果在实验之前不可知。

令 $\Omega = \{\text{相同条件下所有随机实验的结果}\}$ ，称为样本空间，任何 Ω 的子集 A 称为一个事件。

事件有如下运算：

- (1). A^c ，对立事件；
- (2). $A \cup B$ ， A 发生或 B 发生；
- (3). $A \cap B$ ， A 发生且 B 发生，可以简写为 AB ；
- (4). $A \setminus B$ ， A 发生且 B 不发生（如果 $B \subset A$ ，也可以写成 $A - B$ ）；
- (5). $A \triangle B = (A \cup B) \setminus (A \cap B)$ ；
- (6). 在 $\{A_n\}$ 中，“出现了无数次的结果全体”和“只有在有限个集合中不出现的集合全体”分别为

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right)$$
$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \left(\bigcap_{k=n}^{\infty} A_k \right)$$

事件域：样本空间 Ω 的子集族 $\mathcal{F}$ 满足：

- (1). $\phi, \Omega \in \mathcal{F}$ ；
 - (2). $\forall A \in \mathcal{F}, A^c \in \mathcal{F}$ ；
 - (3). 可列个 $A_n \in \mathcal{F}$ ，它们的并集 $\bigcup_{n=1}^{\infty} A_i \in \mathcal{F}$ 。
- 称 $\mathcal{F}$ 是样本空间 Ω 上的事件域。

1.2 概率

定义 1.2.1 古典定义：进行 n 次重复实验，记 $n(A)$ 为事件 A 发生的次数，如果 $\frac{n(A)}{n}$ 随着 n 的增大而稳定于某个数 $P(A)$ ，称之为事件 A 发生的概率。

公理化定义：如果映射 $P: \mathcal{F} \rightarrow R$ 满足

- (1). 规范性： $P(\Omega) = 1$ ；
- (2). 非负性： $\forall A \in \mathcal{F}, P(A) \geq 0$ ；
- (3). 可列可加性：可列个 $A_n \in \mathcal{F}$ ，且它们两两不相容，则

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

则称 P 为概率空间 $(\Omega, \mathcal{F})$ 上的概率测度, $(\Omega, \mathcal{F}, P)$ 称为一个概率空间。

定理 1.2.1 概率测度的性质包括:

1. $P(\phi) = 0$;
2. $P(A^c) = 1 - P(A)$;
3. 有限可加性: 有限个 $A_n \in \mathcal{F}, n = 1, 2, \dots, k$, 则 $P(A_1 \cup A_2 \cup \dots \cup A_k) = P(A_1) + \dots + P(A_k)$;
4. $A \subset B \Rightarrow P(A) \leq P(B)$;
5. $P(A \cup B) = P(A \setminus B) + P(B \setminus A) + P(A \cap B) = P(A) + P(B) - P(A \cap B)$, 推广情况:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n p(A_i) - \sum_{1 \leq i < j \leq n} p(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} p(A_i \cap A_j \cap A_k) + \dots + (-1)^{n+1} P(A_1 \cap \dots \cap A_n)$$

6. 如果 $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$, 称 $\{A_n\}$ 为一列单调递增的事件列, 定义

$$A = \bigcup_{n=1}^{\infty} A_n = \lim_{n \rightarrow \infty} A_n$$

则

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

类似地, 如果 $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$, 称 $\{A_n\}$ 为一列单调递减的事件列, 定义

$$A = \bigcap_{n=1}^{\infty} A_n = \lim_{n \rightarrow \infty} A_n$$

则

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

这两个性质分别称为“下连续性”和“上连续性”。

定理 1.2.2 概率函数的第三条性质: 可列可加性, 可由下连续性 + 有限可加等价替代。

例 1.2.1 甲乙玩扔硬币游戏, 甲扔了 $n+1$ 次, 乙扔了 n 次, 问甲抛出的正面个数比乙多的概率?

假设甲、乙抛出的正面个数分别为 a, b , 反面个数分别为 c, d ,

$$\begin{aligned} P(a > b) &= P(c > d) = P(n+1-a > n-b) = P(a < b+1) = P(a \leq b) \\ &= 1 - P(a > b) = \frac{1}{2} \end{aligned}$$

例 1.2.2 现在有 n 封信和其对应的 n 个信封, 现在随机的把信放入信封, 求至少有一个放对的概率。

记事件 A_k 为信封 k 放对了, 显然

$$\begin{aligned} P(A_i) &= \frac{(n-1)!}{n!} \\ P(A_i \cap A_j) &= \frac{(n-2)!}{n!} \\ &\vdots \\ P(A_{i_1} \cap \cdots \cap A_{i_m}) &= \frac{(n-m)!}{n!} \end{aligned}$$

所求的概率为

$$\begin{aligned} P(A_1 \cup \cdots \cup A_k) &= C_n^1 \frac{(n-1)!}{n!} - C_n^2 \frac{(n-2)!}{n!} + \cdots + (-1)^{n+1} C_n^n \frac{(n-n)!}{n!} \\ &= \sum_{k=1}^n C_n^k \frac{(n-k)!}{n!} (-1)^{k-1} \\ &= \sum_{k=1}^n \frac{(-1)^{k-1}}{k!} \end{aligned}$$

值得一提的是, 当 n 趋于无穷大的时候, 上述概率等于 $1 - e^{-1}$, 也就是说, 全放错的概率恰好是 e^{-1} .

1.3 条件概率

定义 1.3.1 设事件 A 的概率 $P(A) > 0$, A 发生的前提下 B 发生的概率称为条件概率, 定义为

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

所以有

$$P\left(\bigcap_{k=1}^n A_k\right) = P(A_1) P(A_2|A_1) P(A_3|A_1 A_2) \cdots P(A_n|A_1 \cdots A_{n-1})$$

例 1.3.1 n 段绳子, 随机取两个绳头打结, 求打了 n 个结之后正好有 n 个圈的概率 (即每个绳子都自己和自己打结)。

记事件 A_k 为第 k 次打结正好多形成了一个圈,

$$\begin{aligned} P(A_1) &= \frac{1}{2n-1} \\ P(A_2|A_1) &= \frac{1}{2n-3} \\ &\vdots \\ P(A_k|A_1 A_2 \cdots A_{k-1}) &= \frac{1}{2(n-k)+1} \end{aligned}$$

定理 1.3.1 全概率公式: 如果 $\bigcup_{k=1}^n A_k = \Omega$ 且它们两两不相容, 称 $\{A_k\}$ 是 Ω 的一个分割. 若 $\forall P(A_i) > 0$, 则

$$\forall B \in \mathcal{F}, P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

常用

$$P(B) = P(B|A) P(A) + P(B|A^c) P(A^c)$$

定理 1.3.2 贝叶斯公式: $\{A_i\}$ 是 Ω 的一个分割, 则

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

例 1.3.2 单项选择题有四个选项, 只有一个为正确选项. 同学甲会做这道题的概率为 p , 不会时蒙对的概率为 $\frac{1}{4}$, 如果甲做对了这道题, 那么他真的会做这道题的概率为多少?

设事件 A 为甲会做这道题, B 为甲做对了这道题. 则

$$P(A|B) = \frac{P(B|A) P(A)}{P(B|A) P(A) + P(B|A^c) P(A^c)} = \frac{p^2}{p^2 + \frac{1}{4}(1-p)}$$

例 1.3.3 某病的患病率为 p , 误诊率为 q , 误诊是指患病没有被确诊, 或者没得病但被误诊为患病. 问如果一个人被确诊, 那么他真的得了病的概率是多少?

设事件 A 为此人得病, 事件 B 为此人被确诊. 则

$$P(A|B) = \frac{P(B|A) P(A)}{P(B|A) P(A) + P(B|A^c) P(A^c)} = \frac{(1-q)p}{(1-q)p + q(1-p)}$$

1.4 独立性

定义 1.4.1 如果事件 A, B 满足 $P(AB) = P(A)P(B)$, 称其相互独立.

推论 1.4.1 $P(A) = 0$ 或 1 , 则 A 与任何事件独立.

推论 1.4.2 A 与 B 独立, 则 A^c 与 B 、 A 与 B^c 、 A^c 与 B^c 均相互独立.

注 多个事件 $A_i, i = 1, 2, \dots, n$ 相互独立的条件为: 任取 k 个事件 $A_{i_1}, \dots, A_{i_k}, k = 2, \dots, n$, 有

$$P\left(\bigcap_{m=1}^k A_{i_m}\right) = \prod_{m=1}^k P(A_{i_m})$$

定义 1.4.2 条件独立性: $P(C) > 0, P(A \cap B|C) = P(A|C)P(B|C)$, 称 A, B 关于 C 条件独立.

注 相当于 A, B 在考虑概率测度 $P(\cdot|C)$ 时相互独立. 条件独立和独立之间没有关系.

1.5 乘积空间

定理 1.5.1 $\mathcal{F}, \mathcal{G}$ 是 Ω 上的 σ 代数, 则 $\mathcal{F} \cap \mathcal{G}$ 也是 Ω 上的 σ 代数。

注 $\mathcal{F} \cup \mathcal{G}$ 不一定是 σ 代数。

定义 1.5.1 $\mathcal{A}$ 是 Ω 的部分子集组成的集合类, 满足 $\mathcal{A} \subset \mathcal{F}_i$ 的 σ 代数 $\mathcal{F}_i$ 的交集称为由 $\mathcal{A}$ 生成的 σ 代数, 记作 $\sigma(\mathcal{A})$, 即包含 $\mathcal{A}$ 的最小的 σ 代数。

例 1.5.1 R 上的 Borel σ 代数: 所有的左闭右开区间和开区间设为

$$\mathcal{A} = \{(a, b] | a, b \in R, a < b\}$$

$$\mathcal{B} = \{(a, b) | a, b \in R, a < b\}$$

下面证明, 它们生成的 σ 代数相等,

$$\begin{aligned} \bigcap_{n=1}^{\infty} (b - \frac{1}{n}, b] &= \{b\} \in \sigma(\mathcal{A}) \\ (a, b] \setminus \{b\} &= (a, b) \Rightarrow \sigma(\mathcal{B}) \subset \sigma(\mathcal{A}) \\ \bigcup_{n=m}^{\infty} (a, n) &= (a, \infty) \in \sigma(\mathcal{B}), m = \inf\{m \in N | m \geq a\} \\ (a, \infty) \setminus (b, \infty) &= (a, b] \Rightarrow \sigma(\mathcal{A}) \subset \sigma(\mathcal{B}) \end{aligned}$$

所以 $\sigma(\mathcal{A}) = \sigma(\mathcal{B})$. 既然二者等价, 将其统称为 R 上的 Borel σ 代数, 记作 $\mathcal{B}(R)$. $\mathcal{B}(R)$ 中的元素称为 Borel 集。

$\mathcal{B}(R)$ 的一些性质: 此部分是实分析内容, 详见实分析笔记 1.1 节

1. 开集、闭集是 Borel 集;
2. 闭集的可数并、开集的可数交是 Borel 集;
3. 闭集的可数并的可数交、开集的可数交的可数并是 Borel 集。

1.6 一些例子

例 1.6.1 掷硬币, 正面乙给甲一元钱, 反面甲给乙一元钱, 假设初始甲有 k 元钱, 乙有 $N - k$ 元钱, 重复此游戏直至一方钱输光为止, 求最终甲的钱输光了的概率。

解 记 A_k 为甲拥有 k 元钱并最终输光了的概率, 设 $P(A_k) = p_k$. 记 B 为抛出正面的概率, 则

$$p_k = P(A_k | B) P(B) + P(A_k | B^c) P(B^c) = \frac{1}{2} (p_{k+1} + p_{k-1})$$

又有 $p_0 = 1, p_N = 0$, 可知

$$p_k = 1 - \frac{k}{N}$$

例 1.6.2 古典概型: Ω 是有限集且各元素概率相等。

计数: n 个元素中任取 m 个元素有多少种取法?

1. 放回、有序: n^m

2. 不放回、有序: A_n^m
 3. 放回、无序: 相当于把 m 个等价元素放进 n 个标了序号的抽屉里, 使用隔板法, n 个抽屉意味着 $n-1$ 个隔板, 这 $n-1$ 个隔板和 m 个等价元素作全排列,

$$\frac{(n+m-1)!}{(n-1)!m!} = C_{n+m-1}^m$$

4. 不放回、无序: C_n^m
 把 n 个球放入 N ($N \geq n$) 个盒子中, 求下列事件的概率:
 A. 指定的 n 个盒子中各有一球;
 B. 恰好 n 个盒子中各有一球;
 C. 至少有一个盒子中有两个球。

解

$$\#\Omega = C_{n+N-1}^{N-1}, \#A = 1, \#B = C_N^n, \#C = \#\Omega - \#B$$

若球可区分:

$$\#\Omega = N^n, \#A = n!, \#B = n! \cdot C_N^n = A_N^n, \#C = \#\Omega - \#B$$

第二章 随机变量及分布函数

2.1 概念介绍

定义 2.1.1 掷两个骰子, $\Omega = \{(i, j) | i, j = 1, 2, \dots, 6\}$, X 表示两个点数之和, $X(i, j) = i + j$, 则 $X = l$ 是一个事件。

一般定义: $(\Omega, \mathcal{F}, P)$ 是概率空间, 函数 $X: \Omega \rightarrow R$ 满足 $\{\omega | X(\omega) \leq x\} \in \mathcal{F}, \forall x \in R$. 称 X 是随机变量, 简称 *r.v.*

例 2.1.1 盒子中有 n 个黑球, 1 个白球, 逐个取出球, 设 X 为第 X 次取到白球。

$$\Omega = \{e_i = (0, \dots, 0, 1, 0, \dots, 0), \text{ 第 } i \text{ 个为 } 1, \text{ 其余为 } 0 | i = 1, 2, \dots, n+1\}$$

$$P(\{X = k\}) = \begin{cases} \frac{1}{n+1}, & k = 1, 2, \dots, n+1 \\ 0, & k \neq 1, 2, \dots, n+1 \end{cases}$$

关于 X 的一些命题:

- (1). $X: \Omega \rightarrow R$ 是单值函数;
- (2). $\forall x \in R, \{X \leq x\} \in \mathcal{F}$, 则 $\{X > x\}, \{X = x\}, \{a < x \leq b\}, \{a < x < b\} \in \mathcal{F}$.

证明

$$\begin{aligned} \{X > x\} &= \{X \leq x\}^c \\ \{X = x\} &= \{X \leq x\} \setminus \{X < x\} = \{X \leq x\} \setminus \left(\bigcup_{n=1}^{\infty} \{X \leq x - \frac{1}{n}\} \right) \end{aligned}$$

- (3). 若 B 是 Borel 集, 则 $\{\omega | X(\omega) \in B\} \in \mathcal{F}$.
- (4). X, Y 都是随机变量, 则 $X + Y, X \cdot Y$ 也是随机变量; 如果 $g: R \rightarrow R$ 是 Borel 可测函数, 则 $g(X)$ 也是随机变量。

定义 2.1.2 X 是 $(\Omega, \mathcal{F}, P)$ 上的随机变量, 则称 $F(x) = p(X \leq x)$ 为 X 的分布函数, 简称 *c.d.f*

一些性质:

- (1). $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$;
- (2). $x \leq y \Rightarrow F(x) \leq F(y)$;
- (3). $F(x)$ 右连续。
- (4). $p(X = x) = F(x) - \lim_{a \rightarrow x^-} F(a) = F(x) - F(x-0)$.

满足前三条的函数可以认为是分布函数。

2.1.1 离散型随机变量

定义 2.1.3 X 的取值至多可列个, $F(x) = p(X = x)$. 离散型随机变量的分布律形如下表:

X	X_1	X_2	$\cdots$	X_i	$\cdots$
p	p_1	p_2	$\cdots$	p_i	$\cdots$

$$p_1 + \cdots + p_i + \cdots = 1.$$

例 2.1.2 向目标物开炮, 击中的概率为 p , 每次射击结果独立, 设目标物被累计击中 r 次之后被摧毁, 记 X 目标物被摧毁时开炮的次数, 求 X 的分布律。

解 也就是前 $X-1$ 次击中了 $r-1$ 次, 然后又击中了一次:

$$p_x = p(X = x) = F(x) = \begin{cases} 0, & x < r \\ p^r, & x = r \\ C_{x-1}^{r-1} p^r (1-p)^{x-r}, & x > r \end{cases}$$

2.1.2 连续型随机变量

定义 2.1.4 如果存在非负函数 $f(u)$ 使得 $F(x) = \int_{-\infty}^x f(t) dt$, 称 X 是连续性随机变量。 $f(x)$ 被称为 (概率) 密度函数, 简称 $p.d.f$.

一些性质:

- (1). $f(x) \geq 0$;
- (2). $\int_{-\infty}^{\infty} f(x) dx = 1$.

满足这两条的函数可以认为是密度函数。

定义 2.1.5 如果 $F_1(x)$ 是连续型分布函数, $F_2(x)$ 是离散型分布函数, 形如

$$F(x) = \alpha F_1(x) + (1 - \alpha) F_2(x), \alpha \in (0, 1)$$

的函数被称为混合型分布函数。

2.1.3 随机向量

定义 2.1.6 $(X_1, \cdots, X_n)$ 被称为随机向量, 其中 X_i 都是随机变量。

$F(x_1, \cdots, x_n) = P(\forall i, X_i \leq x_i)$ 被称为联合分布函数。

对于二维连续随机向量 (X, Y) , 称 $F_X(x) = P(X \leq x) = \lim_{y \rightarrow +\infty} F(x, y)$ 为随机变量 X 的边缘分布函数。

定义 2.1.7 二维离散型随机向量 (X, Y) 的取值为至多可列个, 分布列中 $p_{ij} = P(X = X_i, Y = Y_j)$. 二维连续型随机变量 (X, Y) 的联合分布可以被写成

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$$

其中 $f(u, v)$ 称为联合密度函数。它同样有边缘分布: $f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy$.

2.2 离散型随机变量

2.2.1 常见离散型分布

1. 均匀分布 $X(\omega) = \{X_1, \dots, X_n\}$, $P(X = X_n) = \frac{1}{n}$.

2. 二项分布 重复某试验 n 次, 其中事件 A 发生的次数为 X , 则 $X(\omega) = \{0, 1, \dots, n\}$, $P(X = k) = C_n^k p^k (1-p)^{n-k}$, 其中 p 是每次试验事件 A 发生概率。 X 满足二项分布记作 $X \sim B(n, p)$.

定理 2.2.1 如果 $\lim_{n \rightarrow \infty} np_n = \lambda$, 则

$$\lim_{n \rightarrow \infty} C_n^k p_n^k (1-p_n)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

3. 几何分布 重复某试验直到事件 A 发生时, 已经进行的试验次数为 X , 则 $X(\omega) = \mathbb{N}_+$, $P(X = k) = (1-p)^{k-1} p$.

4. 负二项分布 重复某试验直到事件 A 发生 r 次时, 已经进行的试验次数为 X , 则 $X(\omega) = \{r, r+1, \dots\}$, $P(X = k) = C_{k-1}^{r-1} p^r (1-p)^{k-r}$. X 满足负二项分布记作 $X \sim f(n, p)$.

5. Poisson 分布

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

记作 $X \sim P(\lambda)$.

2.2.2 独立性

定义 2.2.1 称随机变量 X, Y 独立是指任意 $\{X = x_i\}, \{Y = y_j\}$ 都独立。

例 2.2.1 现在进行随机次数的试验, 成功次数为 X , 失败次数为 Y , 试验次数 $N \sim P(\lambda)$, 则 X 和 Y 独立, 即

$$P(X = n, Y = m) = P(X = n) P(Y = m)$$

这个看起来不独立, 其实是独立的。

2.2.3 数学期望

定义 2.2.2 离散型随机变量 X , 设 $f(x) = P(X = x)$, 则 $f(x)$ 被称为 (概率) 质量函数, 简称 $p.m.f.$.

如果 $\sum x \cdot f(x)$ 绝对收敛, 则其值称为 X 的数学期望, 记作 $E(X)$.

2.2.4 方差

定义 2.2.3 离散型随机变量 X , $E((X - E(X))^2)$ 称为方差, 记作 $var(X)$. 通过计算可知 $var(X) = E(X^2) - [E(x)]^2$.

$\sqrt{var(X)}$ 为标准差。

2.2.5 常见随机分布的期望和方差

2.2.6 示性函数举例

2.2.7 条件分布与条件期望

定义 2.2.4 如果 $E((X - E(X))(Y - E(Y)))$ 存在, 称之为协方差, 记作 $cov(X, Y)$. 通过计算可知 $cov(X, Y) = E(XY) - E(X)E(Y)$, 定义

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)}\sqrt{var(Y)}}$$

为 X, Y 的相关系数。

性质:

- (1). $cov(X, Y) = cov(Y, X)$
 - (2). $cov(X, a) = 0, a \in R$
 - (3). $cov(aX, bY) = ab \cdot cov(X, Y)$
 - (4). $cov(X + Y, Z) = cov(X, Z) + cov(Y, Z)$
 - (5). $var(X + Y) = var(X) + var(Y) + 2cov(X, Y)$
 - (6). $|\rho(X, Y)| \leq 1$, 等号成立当且仅当 $\exists a, b \in R, P(Y = aX + b) = 1$.
- X, Y 相互独立 $\Rightarrow cov(X, Y) = 0$, 反之不成立。

多项分布

条件期望

定义 2.2.5 $\Omega = \bigsqcup_{i=1}^n B_i$, $B_i = \{\omega | Y(\omega) = y_i\}$, 定义

$$f_{X|Y}(x|y) = P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

为在 $Y = y$ 条件下, X 的条件分布列。

$$F_{X|Y}(x|y) = P(X \leq x | Y = y)$$

为条件分布函数,

$$E(x|Y = y) = \sum_x x \cdot f_{X|Y}(x|y)$$

称为 $Y = y$ 时 X 的条件期望, 记 $\psi(y) = E(X|Y = y)$, 于是 $\psi(Y(\omega))$ 是一个随机变量。

例 2.2.2 随即给出 N 个随机变量 X_i , $N \sim P(\lambda)$, $X_i \sim B(1, p)$, N 以及 X_i 互相独立,, 令 $X = \sum_{i=1}^N X_i$, 求 $E(X|N), E(N|X), E(X), var(X)$.

解

$$f_{X|N}(x|n) = P(X = x|N = n) = C_n^x p^x q^{n-x}$$

$$f_{N|X}(n|x) = P(N = n|X = x) = \frac{P(N = n, X = x)}{P(X = x)}$$

$$\begin{aligned} &= \frac{P(X = x|N = n)P(N = n)}{\sum_{m \geq x} P(X = x|N = m)P(N = m)} \\ &= \frac{C_n^x p^x q^{n-x} e^{-\lambda} \lambda^n (n!)^{-1}}{\sum_{m \geq x} C_m^x p^x q^{m-x} e^{-\lambda} \lambda^m (m!)^{-1}} = \cdots = \frac{(\lambda q)^{n-x}}{(n-x)!} e^{-\lambda q} \end{aligned}$$

$$E(X|N = n) = \sum_x x \cdot f_{X|N}(x|n) = \sum_x x \cdot C_n^x p^x q^{n-x} = np \Rightarrow E(X|N) = Np$$

$$E(N|X = x) = \sum_{n \geq x} n \cdot f_{N|X}(n|x) = \sum_{n \geq x} n \cdot \frac{(\lambda q)^{n-x}}{(n-x)!} e^{-\lambda q} \stackrel{k=n-x}{=} \sum_{k=0}^{\infty} (k+x) \cdot \frac{(\lambda q)^k}{k!} e^{-\lambda q} = \lambda q + x$$

$$E(X) = E(E(X|N)) = E(Np) = \lambda p$$

$$E(X^2) = \sum_{n=1}^{\infty} E(X^2|N = n)P(N = n) = \sum_{n=1}^{\infty} E\left(\left(\sum_{i=1}^n X_i\right)^2\right)P(N = n)$$

$$= \sum_{n=1}^{\infty} \left(\text{var}\left(\sum_{i=1}^n X_i\right) + \left(E\left(\sum_{i=1}^n X_i\right)\right)^2 \right) \cdot P(N = n)$$

$$= \sum_{n=1}^{\infty} \left(n \cdot \text{var}(X_1) + (np)^2 \right) \cdot P(N = n)$$

$$= \text{var}(X_i)E(N) + p^2 E(N^2)$$

$$\text{var}(X) = E(X^2) - E(X)^2 = \text{var}(X_1)E(N) + E(X_1)^2 \text{var}(N)$$

2.2.8 母函数

随机变量的和

定义 2.2.6 $X \sim f_X, Y \sim f_Y, Z = X + Y$. 于是

$$f_Z(z) = P(Z = z) = P\left(\bigcup_x (\{X = x\} \cap \{Y = z - x\})\right) = \sum_x P(X = x, Y = z - x)$$

也有 $f_Z(z) = \sum_y P(X = z - y, Y = y)$.

如果 X, Y 独立, 则 $f_Z(z) = \sum_x f_X(x) f_Y(z - x) = f_X * f_Y(z)$, 即卷积。

母函数 (生成函数)

定义 2.2.7 对于数列 $\{a_n\}_{n=0}^{\infty}$, $\sum_{n=0}^{\infty} a_n x^n = G_a(x)$ 称为母函数。

定义两个数列 $\{a_n\}, \{b_n\}$ 的卷积 $c_n \triangleq \sum_{k=0}^n a_k b_{n-k}$, 则

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k} x^k x^{n-k} = \left(\sum_{n=0}^{\infty} a_n x^n \right) \cdot \left(\sum_{n=0}^{\infty} b_n x^n \right)$$

X 取非负整数值, $p_k = P(X = k)$, 则 $G_X(s) = \sum_{k=0}^{\infty} p_k s^k$ 称为 X 的概率母函数。

一些性质:

1. $G_X(s) = E(s^X)$.
2. 收敛半径为 $R \geq 1$.
3. 母函数与分布列一一对应: $p_k = \frac{G^{(k)}(0)}{k!}$, 为 s^k 的系数。
4. 若 $G'_X(1)$ 收敛, 其值为 $E(X)$; 若 $E(X(X-1)\cdots(X-n+1))$ 存在, 则 $G_X^{(n)}(1) = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1)s^{k-n}p_k = E(X(X-1)\cdots(X-n+1))$.

常见分布的母函数

独立随机变量的和

定理 2.2.2 $X_1, \cdots, X_n$ 取非负整数值, 相互独立, 母函数分别为 $G_1(s), \cdots, G_n(s)$, Y 为它们之和, 则 Y 的母函数为 $G_Y(s) = \prod_{i=1}^n G_i(s)$.

例 2.2.3 掷 5 个骰子, 求点数之和为 15 的概率。

解 设 X_i 为第 i 个骰子的点数, $G_i(s) = \frac{1}{6}(s + s^2 + \cdots + s^6)$. 设 Y 为点数之和, 则

$$G_Y(s) = G_1(s)^5 = \frac{s^5}{6^5}(1+s+\cdots+s^5)^5 = \frac{s^5}{6^5}(1-s^6)^5(1-s)^{-5} = \frac{1}{6^5}(s^5 - 5s^{11} + \cdots) \left(\sum_{r=0}^{\infty} C_{5+r-1}^r s^r \right)$$

于是, s^{15} 的系数为

$$\frac{1}{6^5}(C_{14}^{10} - 5C_8^4)$$

注

$$\frac{1}{(1-x)^n} = (1+x+x^2+\cdots)^n = \sum_{r=0}^{\infty} C_{n+r-1}^r x^r$$

如果定义 $C_{-n}^r = (-1)^r C_{n+r-1}^r$, 则可以扩展二项式定理得到:

$$(1-x)^{-n} = \sum_{r=0}^{\infty} C_{-n}^r (-x)^r$$

例 2.2.4 随机个独立同分布 (取非负整数值) 随机变量之和为 Y , 即 $Y = \sum_{i=1}^N X_i$, 求 Y 的母函数。

$$\begin{aligned} G_Y(s) &= \sum_y s^Y P(Y=y) = E(s^Y) = E(E(s^Y|N)) = \sum_n E(s^Y|N=n) P(N=n) \\ &= \sum_n E(s^{X_1+\cdots+X_n}|N=n) P(N=n) = \sum_n E(s^{X_1+\cdots+X_n}) P(N=n) \\ &= \sum_n \left(\prod_{k=1}^n G_{X_k}(s) \right) P(N=n) = \sum_n (G_{X_1}(s))^n P(N=n) = G_N(G_{X_1}(s)) \end{aligned}$$

2.2.9 随机游走

模型介绍

定义 2.2.8 一个人站在一维数轴上, 初始位置为 s_0 , 每个时间单位都有 p 的概率往右走一格, 或者 $1-p$ 的概率往左走一格。经过 n 个时间单位之后, 即走了 n 步之后的位置为 s_n . 这一模型被称为 (一维) 随机游走。

例 2.2.5 无限制: 初始位置 $s_0 = a$, 求 $P(s_n = b)$.

解 假设往左走了 l 步, 往右走了 r 步, 于是

$$\begin{cases} l + r = n \\ r - l = b - a \end{cases} \Rightarrow \begin{cases} r = \frac{n+b-a}{2} \\ l = \frac{n-b+a}{2} \end{cases}$$

每种游走情况称为“轨道”: $rlrrll \dots$. 此题情境下的轨道数量为 C_n^r , 于是

$$P(s_n = b) = C_n^{\frac{n+b-a}{2}} p^{\frac{n+b-a}{2}} (1-p)^{\frac{n-b+a}{2}}$$

例 2.2.6 带吸收壁随机游走: 假设人走到 0 或者 N 点时停止, 初始位置为 k , 求最终停在 0 的概率。

解 设 p_k 为初始位置为 k 时, 人最终停在 0 的概率。于是 $p_0 = 1, p_N = 0$.

$$\begin{aligned} p_k &= p_{k+1} \cdot p + p_{k-1} \cdot (1-p) \\ \Rightarrow (p_{k+1} - p_k) &= \frac{1-p}{p} (p_k - p_{k-1}) \\ \Rightarrow p_k &= \begin{cases} \frac{r^k - r^N}{1-r^N}, r = \frac{1-p}{p} \neq 1 \\ 1 - \frac{k}{N}, r = 1 \end{cases} \end{aligned}$$

记 X_k 为从 k 出发被 0 吸收时的移动次数, 求其期望 $D_k = E(X_k)$.

解 设事件 R 代表第一步向右走, 则

$$\begin{aligned} E(X_k) &= E(X_k|R)P(R) + E(X_k|R^c)P(R^c) \\ E(X_k|R) &= \sum_x x \cdots P(X_k = x|R) = \sum_x ((x-1) + 1)P(x_{k+1} = x_1) = E(X_{k+1}) + 1 \\ E(X_k|R^c) &= E(X_{k-1}) + 1 \\ \Rightarrow E(X_k) &= (E(X_{k+1}) + 1)p + (E(X_{k-1}) + 1)(1-p) \\ \Rightarrow D_k &= pD_{k+1} + (1-p)D_{k-1} + 1 \\ \Rightarrow p((D_{k+1} - D_k) - (D_k - D_{k-1})) - (1-2p)(D_k - D_{k-1}) &= -1 \end{aligned}$$

相当于一个二阶常微分方程 $pD'' - (1-2p)D' = -1$. 最终解得

$$D_k = \begin{cases} \frac{1}{1-2p} \left(k - N \frac{1-r^k}{1-r^N} \right), r = \frac{1-p}{p} \neq 1 \\ kN - k^2, r = 1 \end{cases}$$

(这里存疑, 不会解这个方程)

定理 2.2.3 无限制随机游走的一般性质:

1. 空间齐次性 $P(s_n = j | s_0 = a) = P(s_n = j + b | s_0 = a + b).$

2. 时间齐次性 $P(s_n = j | s_0 = a) = P(s_{m+n} = j | s_m = a).$

3. Markov 性 (无后效性) $P(s_{n+m} = j | s_0, s_1, \dots, s_m) = P(s_{n+m} = j | s_m).$

即, 无论 $s_0, s_1, \dots, s_{m-1}$ 等于多少, 都不对 m 步之后的事件产生影响, m 步之后的事件完全由 s_m 决定。此即为无后效性。

轨道计数

记 $N_n(a, b)$ 为从 a 到 b 的轨道数, $N_n^0(a, b)$ 为从 a 到 b 且经过 0 的轨道数。

引理 2.2.1 如果 $n, b - a$ 同奇偶, 则 $N_n(a, b) = C_n^{\frac{n+b-a}{2}}$, 否则为 0.

引理 2.2.2 (反射定理) $a, b > 0$, 则 $N_n^0(a, b) = N_n(-a, b).$

定理 2.2.4 (投票定理) $b > 0$, n 和 b 同奇偶, 从 0 出发, 到达 b 且中途不经过 0, 轨道数为 $\frac{b}{n} N_n(0, b).$

证明 第一步肯定向右, 所求轨道数为 $N_{n-1}(1, b) - N_{n-1}^0(1, b)$, 展开计算即可。

注 为什么叫投票定理? 此定理来源于以下模型: 甲乙为选举人, 投票者挨个随机投票 (投甲乙概率相等), 最终甲得到了 a 票, 乙得到了 b 票, 且 $a > b$, 问计票过程中, 甲始终领先的概率?

如果令 X 为甲的票数减去乙的票数, 则 X 成为一个随机游走的模型, 且 $s_0 = 0, s_{a+b} = a - b$, 如果甲始终领先, 相当于 X 中途没有经过 0, 轨道数就是 $\frac{a-b}{a+b} N_{a+b}(0, a-b)$, 又知总轨道数为 $N_{a+b}(0, a-b)$, 所以所求概率就是 $\frac{a-b}{a+b}$.

两者能直接相除得到概率的原因是, 游走过程中向前向后概率相等, 因此任何一个同步数轨道的发生概率也相等。

定理 2.2.5 初始位置为 0, 最终到达 b , 且中途不经过 0 的概率为:

$$P(s_1, \dots, s_n \neq 0, s_n = b) = \frac{|b|}{n} P(s_n = b)$$

游走最大值

记 $M_n = \max\{s_i | i = 0, 1, \dots, n\}.$

定理 2.2.6 初始位置为 0, $r \geq 1$, 则

$$P(M_n \geq r, s_n = b) = \begin{cases} P(s_n = b), & b \geq r \\ \left(\frac{q}{p}\right)^{r-b} P(s_n = 2r - b), & b < r \end{cases}$$

推论 2.2.1

$$\begin{aligned}
P(M_n \geq r) &= \sum_{b < r} P(M_n \geq r, s_n = b) + P(s_n \geq r) \\
&= \sum_{b < r} \left(\frac{q}{p}\right)^{r-b} P(s_n = 2r - b) + P(s_n \geq r) \\
&\stackrel{2r-b=c}{=} \sum_{c=r+1}^{\infty} \left(\frac{q}{p}\right)^{c-r} P(s_n = c) + P(s_n \geq r) \\
&= \sum_{c=r+1}^{\infty} \left(\left(\frac{q}{p}\right)^{c-r} + 1 \right) P(s_n = c) + P(s_n = r)
\end{aligned}$$

定理 2.2.7 初始位置为 0, 在 n 时刻首次到达 b 的概率为

$$f_n(b) = \frac{|b|}{n} P(s_n = b)$$

定理 2.2.8 初始位置为 0, $p = \frac{1}{2}$, $T_{2n} = \max\{2k | s_{2k} = 0, k = 1, \dots, n\}$, 则 $P(T_{2n} = 2k) = P(s_{2k} = 0) \cdot P(s_{2n-2k} = 0)$.

反正弦律

根据

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, n \rightarrow \infty$$

于是

$$P(s_{2k} = 0) = C_{2k}^k \left(\frac{1}{2}\right)^{2k} = \frac{(2k)!}{k!k!} \left(\frac{1}{2}\right)^{2k} \sim \frac{\left(\frac{2k}{e}\right)^{2k} \sqrt{4\pi k}}{\left(\frac{k}{e}\right)^{2k} 2\pi k} \left(\frac{1}{2}\right)^{2k} = \frac{1}{\sqrt{\pi k}}, k \rightarrow \infty.$$

$$\begin{aligned}
P\left(\frac{T_{2n}}{2n} \leq x\right) &\sim \sum_{k \leq nx} \frac{1}{\pi \sqrt{k(n-k)}} \\
&= \sum_{k \leq nx} \frac{1}{n} \frac{1}{\pi \sqrt{\frac{k}{n} \left(1 - \frac{k}{n}\right)}} \sim \int_0^x \frac{1}{\pi \sqrt{u(1-u)}} du = \frac{2}{\pi} \arcsin \sqrt{x}, n \rightarrow \infty
\end{aligned}$$

2.3 连续性随机变量

2.3.1 密度函数的独立性

定义 2.3.1 $F_x(x) = p(x \leq X) = \int_{-\infty}^x f(t)dt, f(t) \geq 0, \int_{-\infty}^{+\infty} = 1$. 则 $f(x)$ 被称为概率密度函数, 简称 p.d.f.

那么有:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

$X_1, \dots, X_n$ 是连续性随机变量, 如果

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^n P(X_i \leq x_i) \forall x_i \in \mathbb{R} \cup \{-\infty, +\infty\}$$

称它们独立。 $(X_1, \dots, X_n)$ 的联合密度函数为 $f(x_1, \dots, x_n) = \prod_{i=1}^n f_i(x_i)$.

推论 2.3.1 $X_1, \dots, X_n$ 独立当且仅当

$$f(X_1, \dots, X_n) = \prod_{i=1}^n f_i(x_i)$$

定理 2.3.1 $g_1, \dots, g_n$ 是一元 Borel 可测函数¹, $X_1, \dots, X_n$ 独立, 则 $g(X_1), \dots, g(X_n)$ 独立。

2.3.2 数学期望

定义 2.3.2 若 $\int_{-\infty}^{+\infty} f(x)|x|dx$ 收敛, 记 $E[x] = \int_{-\infty}^{+\infty} xf(x)dx$.

引理 2.3.1 X 为非负连续型随机变量, $E(X)$ 存在, 则

$$E(X) = \int_{-\infty}^{+\infty} P(X > x)dx = \int_0^{+\infty} (1 - F(x))dx$$

一般

$$E(X) = \int_0^{\infty} (1 - F(x))dx - \int_0^{+\infty} F(-x)dx$$

证明 X 的 p.d.f. 记作 $f(x)$, 于是

$$\int_0^{\infty} P(X > x)dx = \int_0^{+\infty} \int_x^{+\infty} f(t)dt dx = \int_0^{+\infty} dt \int_0^t f(t)dx = \int_0^{+\infty} tf(t)dt$$

一般

$$\int_0^{+\infty} F(-x)dx = \int_0^{+\infty} \left(\int_{-\infty}^{-x} f(t)dt \right) dx = \int_{-\infty}^0 dt \int_0^{-t} f(t)dx = \int_{-\infty}^0 -tf(t)dt$$

所以

$$\int_0^{\infty} (1 - F(x))dx - \int_0^{+\infty} F(-x)dx = \int_{-\infty}^{\infty} tf(t)dt = E[x]$$

定理 2.3.2 $X, g(X)$ 都是连续型随机变量, X 的 p.d.f. 是 $f(x)$, $\int_{-\infty}^{+\infty} |g(x)|f(x)dx < +\infty$. 则

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

¹就是指 $x : g(x) > a$ 是 Borel 集, $\forall a$.

证明

$$\begin{aligned}
 E(g(X)) &= \int_0^{+\infty} P(g(X) > y) dy - \int_0^{+\infty} P(g(X) < -y) dy \\
 &= \int_0^{+\infty} \int_{\{g(x) > y\}} f(x) dx dy - \int_0^{+\infty} \int_{\{g(x) < -y\}} f(x) dx dy \\
 &= \int_{\{g(x) > 0\}} dx \int_0^{g(x)} f(x) dy - \int_{\{g(x) < 0\}} dx \int_0^{-g(x)} f(x) dy \\
 &= \int_{-\infty}^{+\infty} g(x) f(x) dx
 \end{aligned}$$

定义 2.3.3 $E(X^k) = \int_{-\infty}^{+\infty} x^k f(x) dx$ (绝对收敛时) 称为 X 的 k 阶矩;
 $E((X - E(X))^k)$ 称为 X 的 k 阶中心矩;
 $var(X) = E[(X - E(X))^2] = E(X^2) - (E(X))^2$.

2.3.3 常用连续型分布

均匀分布

$$X \sim u([a, b])$$

$$\begin{aligned}
 f(x) &= \begin{cases} \frac{1}{b-a} & , x \in [a, b] \\ 0 & , x \notin [a, b] \end{cases} \quad F(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , x \in [a, b] \\ 1 & , x \geq b \end{cases} \\
 E[X] &= \frac{a+b}{2} \\
 var(X) &= \frac{(b-a)^2}{12}
 \end{aligned}$$

指数分布

$$X \sim \text{Exp}(\lambda)$$

$$\begin{aligned}
 f(x) &= \begin{cases} \lambda e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases} \\
 E[X] &= \frac{1}{\lambda} \\
 var(X) &= \frac{1}{\lambda^2}
 \end{aligned}$$

例 2.3.1 背景：电子元件寿命相互独立，发生故障事件的间隔。

记 $N(t)$ 为 $[0, t]$ 时段内的事件数， $N(t), N(t+s) - N(s)$ 独立，

$$P(N(h) = 1) \approx \lambda h, N(t) \sim P(\lambda t)$$

X 表示第一个粒子观测到的时刻：

$$P(x \leq t) = P(N(t) \geq 1) = 1 - P(N(t) = 0) = 1 - e^{-\lambda t} \Rightarrow f_X(t) = \lambda e^{-\lambda t}$$

分布函数为

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

进一步得到 $E[X] = \frac{1}{\lambda}$, $E[X^2] = \frac{2}{\lambda^2}$, $\text{var}(X) = \frac{1}{\lambda^2}$.

定理 2.3.3 (无记忆性) X 为非负连续型随机变量, X 服从指数分布当且仅当 $P(X > s + t | X > t) = P(X > s), t, s > 0$.

正态分布

$$X \sim N(\mu, \sigma^2)$$

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ E(X) &= \mu \\ \text{var}(X) &= \sigma^2 \end{aligned}$$

我们把 $N(0, 1)$ 叫做标准正态分布, 其密度函数

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

分布函数 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$.

推论 2.3.2 $X \sim N(\mu, \sigma^2)$, 则 $aX + b \sim N(a\mu + b, a^2\sigma^2)$.
特别地, $\frac{X-\mu}{\sigma} \sim N(0, 1)$.

Γ 分布

$$X \sim \Gamma(\alpha, \lambda), \alpha, \lambda > 0$$

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

例 2.3.2 背景: 第 n 个独立事件发生时间 $T_n \sim \Gamma(n, \lambda)$, $N(X) \sim P(\lambda X)$.

$$P(T_n \leq X) = P(N(X) \geq n) = 1 - \sum_{k=0}^{n-1} P(N(X) = k) = 1 - \sum_{k=0}^{n-1} e^{-\lambda X} \cdot \frac{(\lambda X)^k}{k!}$$

上式求导就得到 Γ 分布的密度函数。

$$E(X) = \frac{\lambda}{a}, \text{var}(X) = \frac{\alpha}{\lambda^2}$$

Beta 分布

$$X \sim B(a, b), a, b > 0$$

$$f(x) = \begin{cases} \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} & , x \in (0, 1) \\ 0 & , x \notin (0, 1) \end{cases}$$

$$E(x) = \frac{a}{a+b}, E(X^2) = \frac{(a+1)a}{(a+b+1)(a+b)}$$

Cauchy 分布

$$f(x) = \frac{1}{\pi(1+x^2)}$$

$$F(x) = \frac{\arctan x + \frac{\pi}{2}}{\pi}$$

期望不存在。

2.3.4 连续型随机向量**期望、方差与协方差**

定义 2.3.4 $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$.

边缘分布: $P(X \leq x) = \int_{-\infty}^x du \int_{-\infty}^{+\infty} f(u, v) dv$

X 概率密度: $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$.

随机变量 $g(X, Y)$ 的期望:

$$E(g(X, Y)) = \iint_{\mathbb{R}^2} g(x, y) f(x, y) dx dy.$$

单个随机变量的期望、方差:

$$E[X] = \iint_{\mathbb{R}^2} x f(x, y) dx dy$$

$$var[X] = \iint_{\mathbb{R}^2} (x - E(X))^2 f(x, y) dx dy$$

协方差:

$$covar(X, Y) = E[XY] - E[X]E[Y]$$

相关系数:

$$\rho = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}}$$

例 2.3.3 (X, Y) 在区域 $D = \{(x, y) : x^2 + y^2 \leq R^2\}$ 上均匀分布,

$$f(x, y) = \begin{cases} \frac{1}{\pi R^2} & , (x, y) \in D \\ 0 & , (x, y) \notin D \end{cases}$$

1° 求边缘分布的概率密度

2° $\rho = \sqrt{X^2 + Y^2}$, 求 $E(\rho)$.

解

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \frac{2}{\pi R^2} \sqrt{R^2 - x^2}, -R \leq x \leq R$$

$$f_Y(y) = \frac{2}{\pi R^2} \sqrt{R^2 - y^2}, -R \leq y \leq R$$

$P(\rho \leq x) = \frac{x^2}{R^2}$ 是分布函数, 于是密度函数等于 $f_\rho(x) = \frac{2x}{R^2}$. 进而期望 $E(\rho) = \int_0^R x \frac{2x}{R^2} dx = \frac{2}{3}R$.

定理 2.3.4 $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ 是 Borel 可测函数, (X, Y) 是连续型随机变量, 则 $g(X, Y)$ 是连续性随机变量, 且期望 (如果存在) 为

$$E(g(X, Y)) = \iint_{\mathbb{R}^2} g(x, y) f(x, y) dx dy$$

例 2.3.4 (二元正态分布相关) X, Y 是连续型随机变量, 称它们服从二元正态分布 $(X, Y) \sim N(0, 1; 0, 1; \rho)$, 当联合概率密度函数

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}$$

于是

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y] = \rho$$

一般情况: $(X, Y) \sim N(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2; \rho)$

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}Q(x, y)}$$

其中

$$Q(x, y) = \frac{(x - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x - \mu_1)(y - \mu_2)}{\sigma_1\sigma_2} + \frac{(y - \mu_2)^2}{\sigma_2^2}$$

$$1^\circ f_X(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$$

$$2^\circ \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} = \rho$$

$$3^\circ \rho = 0 \text{ 时意味着 } X, Y \text{ 独立。}$$

连续型随机变量的条件分布

定义 2.3.5 $f_Y(y) > 0, f_Y(y)\Delta y > 0$

$$\begin{aligned}\lim_{\Delta y \rightarrow 0} p(X \leq x | y \leq Y \leq y + \Delta y) &= \lim_{\Delta y \rightarrow 0} \frac{p(X \leq x, y \leq Y \leq y + \Delta y)}{p(y \leq Y \leq y + \Delta y)} \\ &= \lim_{\Delta y \rightarrow 0} \frac{\int_{-\infty}^x du \int_y^{y+\Delta y} f(u, y) dv}{\int_{-\infty}^{+\infty} f(u, y) du} = \int_{-\infty}^x \frac{f(u, y)}{f_Y(y)} du\end{aligned}$$

由此引出定义：给定 $Y = y$ 条件下， X 的条件密度 $f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, (f_Y(y) > 0)$

$$F_{X|Y}(x|y) = \int_{-\infty}^x f_{X|Y}(x|y) du, \text{ 给定 } Y = y \text{ 条件下, } X \text{ 的条件分布函数}$$

条件期望：

$$\psi(x) = E[Y|X = x] = \int_{-\infty}^{+\infty} y f_{Y|X}(y|x) dy$$

例 2.3.5 $(X, Y) \sim N(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$, 求 $f_{X|Y}(x|y)$.

解

$$\begin{aligned}f_{X|Y}(x|y) &= \frac{f(x, y)}{f_Y(y)} \\ &= \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right) \right]}{\frac{1}{\sqrt{2\pi}\sigma_2} \exp \left(-\frac{(y-\mu_2)^2}{2\sigma_2^2} \right)} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp \left(-\frac{1}{2(1-\rho^2)\sigma_1^2} \left((x-\mu_1) - \rho \frac{\sigma_1}{\sigma_2} (y-\mu_2) \right)^2 \right)\end{aligned}$$

在给定 $Y = y$ 条件下， $X \sim N(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2), \sigma_1^2(1 - \rho^2))$

例 2.3.6 $(X, Y) \sim u(D), D = \{x^2 + y^2 \leq r^2\}$, 密度函数为

$$f(x, y) = \frac{1}{\pi r^2}, (x, y) \in D$$

求 $f_{X|Y}(x|y)$.

$$\text{解 } f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \frac{1}{\pi r^2} dx = \frac{2}{\pi r^2} \sqrt{r^2-y^2}, -r \leq y \leq r.$$

$$f_{X|Y}(x|y) = \frac{1}{\pi r^2} \frac{\pi r^2}{2\sqrt{r^2-y^2}} = \frac{1}{2\sqrt{r^2-y^2}}, -\sqrt{r^2-y^2} \leq x \leq \sqrt{r^2-y^2}$$

$Y = y$ 条件下， X 服从 $[-\sqrt{r^2-y^2}, \sqrt{r^2-y^2}]$ 上的均匀分布。

例 2.3.7 (X, Y) 联合密度

$$f(x, y) = \begin{cases} \frac{1}{x}, & 0 < y \leq x < 1 \\ 0, & \text{otherwise} \end{cases}$$

求 $f_{Y|X}(y|x)$ 和 $P(X^2 + Y^2 \leq 1 | X = x)$.

解 $0 < x < 1$ 时, $f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_0^x \frac{1}{x}dy = 1$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{1}{x}, 0 < y < x$$

$$\begin{aligned} P(X^2 + Y^2 \leq 1 | X = x) &= P(-\sqrt{1 - X^2} \leq Y \leq \sqrt{1 - X^2} | X = x) \\ &= \begin{cases} P(0 \leq Y \leq X | X = x) & , 0 < x \leq \frac{\sqrt{2}}{2} \\ P(0 \leq Y \leq \sqrt{1 - x^2} | X = x) & \frac{\sqrt{2}}{2} < x < 1 \end{cases} \\ &= \begin{cases} \int_0^x \frac{1}{x}dy = 1 & , 0 < x \leq \frac{\sqrt{2}}{2} \\ \int_0^{\sqrt{1-x^2}} \frac{1}{x}dy = \frac{\sqrt{1-x^2}}{x} & , \frac{\sqrt{2}}{2} < x < 1 \end{cases} \end{aligned}$$

例 2.3.8 设 $X_1, \dots, X_n$ 独立, 且均服从 $(0, 1)$ 上的均匀分布, 设

$$N = \min\{n | \sum_{i=1}^n X_i > 1\}$$

求 $E[N]$.

(0 到 1 随机取数字, 平均取几个和才能超过 1?)

解 设 $N(x) = \min\{n | \sum_{i=1}^n X_i > x\}$, $m(x) = E(N(x)) = E(E(N(x)|x_1))$

$$\begin{aligned} E(N(x)|x_1 = y) &= \begin{cases} 1 & , y > x \\ 1 + m(x - y) & , y \leq x \end{cases} \\ m(x) &= E(E(N(x)|x_1)) = \int_0^1 E(N(x)|x_1 = y)f_{x_1}(y)dy \\ &= \int_0^x 1 + m(x - y)dy + \int_x^1 1dy = 1 + \int_0^x m(t)dt. \end{aligned}$$

解得 $m(x) = e^x$, 进而 $E[N] = m(1) = e$.

2.4 随机变量的函数

定理 2.4.1 (单变量) X 的概率密度为 $f(x)$, $Y = g(x) \in C^1$ 且严格单调, 则 Y 的概率密度

$$f_Y(y) = f(g^{-1}(y)) \cdot |g^{-1}(y)'|$$

证明

$$\begin{aligned} P(Y \leq y) &= P(x \in \{g(x) \leq y\}) = \int_{\{g(x) \leq y\}} f(x) dx \\ &= \int_{-\infty}^y f(g^{-1}(t)) |g^{-1}(t)'| dt \\ \Rightarrow f_Y(y) &= f(g^{-1}(y)) \cdot |g^{-1}(y)'| \end{aligned}$$

推论 2.4.1 如果 $g(x)$ 在不重叠的区间段 $I_1, \dots, I_n$ 上分别严格单调, I_i 上确定反函数 $x = h_i(y)$ 有连续导数, 那么

$$f_Y(y) = \sum_i f(h_i(y)) \cdot |h_i(y)'|$$

例 2.4.1 X 的分布函数 $F(x)$ 严格单调递增且连续, 则 $Y = F(X) \sim u([0, 1])$.

证明

$$P(F(x) \leq y) = P(X \leq F^{-1}(y)) = \begin{cases} 0 & , y < 0 \\ F(F^{-1}(y)) = y & , 0 \leq y < 1 \\ 1 & , y \geq 1 \end{cases}$$

注 如果 $\theta \sim u([0, 1])$, 对于任意严格单调递增分布函数 $F(x)$, 可以定义一个随机变量服从 $F(x)$: $X = F^{-1}(\theta)$.

定理 2.4.2 (多变量) 如果 $Y_1 = g_1(X_1, X_2), Y_2 = g_2(X_1, X_2)$ 满足:

(1) 可以确定逆映射:

$$\begin{cases} x_1 = h_1(y_1, y_2) \\ x_2 = h_2(y_1, y_2) \end{cases}$$

(2) 雅可比行列式非 0, 即

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} \neq 0$$

则 (Y_1, Y_2) 有联合密度 $f_Y(y_1, y_2) = f(h_1(y_1, y_2), h_2(y_1, y_2)) \cdot |J|^{-1}$.

例 2.4.2 X, Y 相互独立, 都服从 $N(0, 1)$, 令

$$R = \sqrt{X^2 + Y^2}, \Theta = \begin{cases} \arctan \frac{Y}{X} & , X > 0, Y > 0 \\ \pi + \arctan \frac{Y}{X} & , X < 0 \\ 2\pi + \arctan \frac{Y}{X} & , X > 0, Y < 0 \end{cases}$$

求 R, Θ 的分布。

解 $f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}, \frac{\partial(x, y)}{\partial(r, \theta)} = r$, 于是

$$f_{R, \Theta}(r, \theta) = \frac{1}{2\pi} r e^{-\frac{r^2}{2}}$$

R, Θ 独立, $\Theta \sim u([0, 2\pi]) = 2\pi u_1$, $f_R(r) = re^{-\frac{r^2}{2}}$

$$F_R(r) = 1 - e^{-\frac{r^2}{2}} = u_2 \Rightarrow r = \sqrt{-2\ln(1-u_2)}$$

实际上, $u_1, u_2 \sim u([0, 1])$ 且相互独立, 此时

$$\begin{aligned} \Theta &= 2\pi u_1 & X &= \sqrt{-2\ln u_2} \cos(2\pi u_1) \\ R &= \sqrt{-2\ln u_2} & Y &= \sqrt{-2\ln u_2} \sin(2\pi u_1) \end{aligned}$$

例 2.4.3 $(X, Y) \sim f(x, y)$, 求 $Z = X + Y$ 的分布。

解 以前的方法:

$$P(X + Y \leq a) = \iint_{x+y \leq a} f(x, y) dx dy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{a-x} f(x, y) dy$$

现在的方法: 写出 $(X + Y, X)$ 的联合分布, 然后求 $X + Y$ 的边缘分布。

例 2.4.4 (顺序统计量) n 个随机变量 $X_1, \dots, X_n$ 独立同分布, 分布函数都是 $F(x)$. 把它们从小到大排列, 排在第 k 位的记作 X_k^* , 求其概率密度。

解 $\omega \in (X_k^* \leq x) \Leftrightarrow X_1(\omega), \dots, X_n(\omega)$ 中至少有 k 个 $\leq x$. 设

$$A_m = \{\omega | X_1(\omega), \dots, X_n(\omega) \text{ 中恰好有 } m \text{ 个 } \leq x\}$$

则

$$\begin{aligned} P(A_m) &= C_n^m (F(x))^m (1 - F(x))^{n-m} \\ P(X_k^* \leq x) &= \sum_{i=k}^n P(A_i) = \sum_{i=k}^n C_n^i (F(x))^i (1 - F(x))^{n-i} \\ f_k(x) &= \frac{d}{dx} \left(\sum_{i=k}^n C_n^i (F(x))^i (1 - F(x))^{n-i} \right) \\ &= \frac{n!}{(k-1)!(n-k)!} (F(x))^{k-1} (1 - F(x))^{n-k} f(x) \end{aligned}$$

2.5 多元正态分布

我们前面已经介绍过二元正态分布: $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left(\left(\frac{x_1-\mu_1}{\sigma_1} \right)^2 - 2\rho \frac{x_1-\mu_1}{\sigma_1} \frac{x_2-\mu_2}{\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2} \right)^2 \right)}$$

设 $\vec{x} = (x_1, x_2)$, $\vec{\mu} = (\mu_1, \mu_2)$,

$$\begin{aligned} \Sigma &= \begin{pmatrix} \text{cov}(x_1, x_1) & \text{cov}(x_1, x_2) \\ \text{cov}(x_2, x_1) & \text{cov}(x_2, x_2) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \\ |\Sigma| &= \sigma_1^2\sigma_2^2(1-\rho^2) > 0 \\ \Sigma^{-1} &= \frac{1}{1-\rho^2} \begin{pmatrix} \sigma_1^{-2} & -\rho(\sigma_1\sigma_2)^{-1} \\ -\rho(\sigma_1\sigma_2)^{-1} & \sigma_2^{-2} \end{pmatrix} \end{aligned}$$

于是得到

$$f(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T}$$

推广到多元:

定义 2.5.1 $\vec{x} = (x_1, \dots, x_n)$, $\vec{\mu} = (\mu_1, \dots, \mu_n)$, Σ 是 n 阶正定对称矩阵, 那么

$$f(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T}$$

如果 $(Y_1, \dots, Y_n)$ 服从 n 元标准正态², $\vec{X} = \vec{Y} \cdot A + \vec{\mu}$ 服从参数为 $\vec{\mu}, \Sigma = A^T A$ 的 n 元正态分布, 记作 $\vec{X} \sim N(\vec{\mu}, \Sigma)$ 。

定理 2.5.1 $\vec{X} \sim N(\vec{\mu}, \Sigma)$, 则:

1° $E[\vec{X}] = \vec{\mu}$

2° Σ 就是协方差矩阵.

3° $\vec{X}$ 的各分量相互独立当且仅当 $\text{cov}(X_i, X_j) = \begin{cases} 0 & , i \neq j \\ \Sigma_{ii} & , i = j \end{cases}$

定理 2.5.2 $\vec{X} \sim N(\vec{\mu}, \Sigma)$, $\vec{Y} = \vec{X} \cdot D \sim N(\vec{\mu}D, D^T \Sigma D)$.

定理 2.5.3 $\vec{X} \sim N(\vec{\mu}, \Sigma)$, 如果 Σ 是准对角矩阵 $\text{diag}(\Sigma_{11}, \Sigma_{22})$, 分别为 l 、 $n-l$ 阶矩阵, 则 $\vec{X}_{11} = (X_1, \dots, X_l) \sim N(\mu_{11}, \Sigma_{11})$.

定理 2.5.4 $\vec{X} \sim N(\vec{\mu}, \Sigma)$, Σ 分块左上角方阵为 Σ_{11} , 那么 $\vec{X}_{11} \sim N(\mu_{11}, \Sigma_{11})$.

定理 2.5.5 $\vec{X} \sim N(\vec{\mu}, \Sigma)$, $A_{n \times m} (n > m)$ 满秩, 则 $\vec{Y} = \vec{X}A \sim N(\vec{\mu}A, A^T \Sigma A)$

χ^2 分布

定义 2.5.2 $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ 相互独立,

记 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ 为样本均值, $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 为样本方差, 则 $E(\bar{X}) = \mu$, $E(s^2) = \sigma^2$.

密度函数 $f(x) = \frac{1}{\Gamma(\frac{d}{2}) 2^{\frac{d}{2}}} x^{\frac{d}{2}-1} e^{-\frac{x}{2}}, x > 0$. 称之为自由度为 d 的 χ^2 的分布, 记作 $\chi^2(d)$.

引理 2.5.1 $Y_1, \dots, Y_n \sim N(0, 1)$, 相互独立, 则 $X = \sum Y_i^2 \sim \chi^2(n)$.

定理 2.5.6 $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ 相互独立, 则

1° $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

2° $\bar{X}$ 与 s^2 相互独立

3° $s^2 \cdot \frac{n-1}{\sigma^2} \sim \chi^2(n-1)$

²指相互独立且服从 $N(0, 1)$.

第三章 特征函数及其应用

3.1 数学期望

3.1.1 抽象积分

定义 3.1.1

1° 简单随机变量 (只取有限个值): 定义 $E(X) = \sum_{i=1}^n X_i P(A_i)$.

2° 非负随机变量 X : $A_n = \{\omega | X > n\}$, $A_{n_i} = \{\omega | \frac{i-1}{2^n} \leq x < \frac{i}{2^n}\}$,

$$X_n = \sum_{i=1}^{n \cdot 2^n} \frac{i-1}{2^n} I_{n_i} + n \cdot I_{A_n}, X_n \nearrow X$$

定义 $E(X) = \lim_{n \rightarrow \infty} E(X_n)$.

3° 一般随机变量 X : 分成 $X = X^+ - X^-$.

记作

$$\int_{\Omega} X(\omega) dp = E(X)$$

推论 3.1.1 $E[g(X)] = \int_{\Omega} g(X) dp = \int_{\Omega} g(x) dF(x)$

$X_n(\omega) \rightarrow X(\omega), \forall \omega \in \Omega, P(\Omega) = 1$.

1° 单调收敛: $0 \leq X_n \leq X_{n+1}$

2° 控制收敛: $\exists$ r.v. Y s.t. $|X_n| \leq Y, E[Y] < \infty$.

3° 有界收敛: $|X_n| \leq c$.

定理 3.1.1 X, Y 相互独立, 期望收敛, $E(XY) = E(X)E(Y)$.

定理 3.1.2 X, Y 同分布 $\Leftrightarrow \forall$ 有界连续函数 $g, E[g(X)] = E[g(Y)]$.

定理 3.1.3 $k > 0, E(|X|^k) < \infty$, 则对 $\forall 0 < r < k, E(|x|^r) < \infty$ 且 $(E|X|^r)^{\frac{1}{r}} \leq (E|X|^k)^{\frac{1}{k}}$

推论 3.1.2 (Jensen inequality) $g(x)$ 是凸函数, 则 $E[g(x)] \geq g(E[x])$.

3.2 特征函数

定义 3.2.1 我们前面已经介绍过离散型随机变量的母函数 $G[S^X] = \sum P(X = j)s^j$.

推广：矩母函数

$$\begin{aligned} M(t) &= E[e^{tx}] \\ &= E\left[1 + tx + \frac{(tx)^2}{2!} + \cdots\right] = \sum_{n=0}^{\infty} \frac{E(X^n)}{n!} t^n \\ E(X^n) &= M^{(n)}(0) \end{aligned}$$

性质：若 $M(t) < \infty$ (当 $|t| < r$)，则

- 1° $E[X^k] = M^{(k)}(0)$.
- 2° X, Y 相互独立，则 $M_{X+Y} = M(X)M(Y)$.
- 3° $M_X(t) = M_Y(t)$ ，则 X, Y 同分布。

例 3.2.1 $X \sim N(0, 1)$,

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tX} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2} dx = e^{\frac{t^2}{2}}$$

$Y \sim N(\mu, \sigma^2), Y = \sigma X + \mu$

$$M_Y(t) = e^{t\mu + \frac{1}{2}t^2\sigma^2}$$

联合矩母函数 $(X_1, \dots, X_n)$, $M(t_1, \dots, t_n) = E[e^{t_1X_1 + \cdots + t_nX_n}]$.

如果相互独立，则 $M(t_1, \dots, t_n) = \prod_{i=1}^n M_{X_i}(t_i)$.

定义 3.2.2 矩母函数的缺点是不一定存在，例如 $f(x) = \frac{1}{(1+x^2)\pi}$.

于是我们引入特征函数： $\varphi(t) = E[e^{itX}] = E[\cos(tX)] + iE[\sin(tX)]$ 一定存在。

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x)$$

特征函数简称 c.f.

定理 3.2.1 特征函数满足：

- 1° $\varphi(0) = 1, |\varphi(t)| \leq 1, \varphi(-t) = \overline{\varphi(t)}$
- 2° $\varphi(t)$ 一致连续
- 3° $\varphi(t)$ 非负定，即：对于 $\forall t_1, \dots, t_n \in \mathbb{R}, z_1, \dots, z_n \in \mathbb{C}$ ，有

$$\sum_{k,j=1}^n \varphi(t_k - t_j) z_k \cdot \bar{z}_j \geq 0$$

定理 3.2.2 如果 $E(|X|^k) < \infty$ ，则 $\varphi^{(j)}(0) = i^j E[X^j], j \leq k$.

$$\varphi(t) = 1 + (it)E[X] + \frac{(it)^2}{2!}E[X^2] + \cdots + \frac{(it)^k}{k!}E[X^k] + o(t^k)$$

定理 3.2.3 X_1, X_2 相互独立, 则 $\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)$.

例 3.2.2 $\varphi(t)$ 是随机变量 X 的特征函数, 则 $\varphi^n(t)$ 也是一个特征函数, $|\varphi(t)|^2 = \varphi(t) \cdot \overline{\varphi(t)}$ 也是特征函数。

如果 $a_n \geq 0, \sum_{n=1}^{\infty} a_n = 1, \{\varphi_n(t)\}$ 是一列特征函数, 则 $\sum_{n=1}^{\infty} a_n \varphi_n(t)$ 也是特征函数。

$$\sum_{n=1}^{\infty} a_n \varphi_n(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{+\infty} e^{itx} d(a_n F_n(x))$$

例 3.2.3 $\varphi(t)$ 是特征函数, 则 $1 - |\varphi(2t)|^2 \leq 4(1 - |\varphi(t)|^2)$

证明 $\operatorname{Re}(1 - \varphi(t)) = \int_{-\infty}^{+\infty} 1 - \cos tx dF(x)$

$$1 - \cos tx = 2\sin^2 \frac{tx}{2} \geq 2\sin^2 \frac{tx}{2} \cos^2 \frac{tx}{2} = \frac{1}{2} \sin^2 tx = \frac{1}{4} (1 - \cos(2tx))$$

所以

$$\operatorname{Re}(1 - \varphi(2t)) \leq 4\operatorname{Re}(1 - \varphi(t))$$

$|\varphi(t)|^2$ 也是特征函数, 故

$$1 - |\varphi(2t)|^2 \leq 4(1 - |\varphi(t)|^2)$$

3.2.1 常见分布的特征函数

Bernoulli 分布

$$\varphi(t) = E[e^{itX}] = 1 - p + p \cdot e^{it}$$

二项分布

$$B(n, p)$$

$$\varphi(t) = (1 - p + p \cdot e^{it})^n$$

指数分布

$$f(x) = \lambda e^{-\lambda x}, x > 0$$

$$\varphi(t) = \frac{\lambda}{\lambda - it}$$

正态分布

$X \sim N(0, 1)$ 时, $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

$$\begin{aligned}\varphi(t) &= \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{+\infty} (\cos tx + i \sin tx) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^{+\infty} \cos tx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ \varphi'(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (-x \sin tx) e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sin tx dx e^{-\frac{x^2}{2}} \\ &= -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t \cos tx e^{-\frac{x^2}{2}} dx = -t \varphi(t)\end{aligned}$$

最终解得 $\varphi(t) = e^{-\frac{t^2}{2}}$.

$$Y = \sigma X + \mu \sim N(\mu, \sigma^2)$$

$$Y_y(t) = E[e^{itY}] = E[e^{it(\sigma X + \mu)}] = e^{it\mu} \varphi_X(\sigma t) = e^{it\mu - \frac{(\sigma t)^2}{2}}$$

3.2.2 反转公式与唯一性定理

定理 3.2.4 (反转公式) X 的分布函数是 $F(x)$, 特征函数为 $\varphi(t)$. 则对于

$$\frac{F(b) + F(b-0)}{2} - \frac{F(a) + F(a-0)}{2} = \lim_{\tau \rightarrow +\infty} \frac{1}{2\pi} \int_{-\tau}^{\tau} \frac{e^{-iat} - e^{-ibt}}{it} \varphi(t) dt$$

定理 3.2.5 (唯一性定理) 分布函数由特征函数唯一确定。

定理 3.2.6 若特征函数 $\varphi(t)$ 满足 $\int_{-\infty}^{\infty} |\varphi(t)| dt < \infty$, 则 $\varphi(t)$ 对应的分布函数 $F(x)$ 可导。

3.2.3 多元特征函数

定义 3.2.3 $(X_1, \dots, X_n)$ 联合分布为 $F(X_1, \dots, X_n)$, 定义

$$\varphi(t_1, \dots, t_n) = E[e^{i(t_1 X_1 + \dots + t_n X_n)}]$$

则有

$$\begin{aligned}\varphi(0, \dots, 0) &= 1, |\varphi(t_1, \dots, t_n)| \leq 1 \\ \varphi(-t_1, \dots, -t_n) &= \overline{\varphi(t_1, \dots, t_n)}\end{aligned}$$

定理 3.2.7 $\varphi(t_1, \dots, t_n)$ 是 $\vec{X} = (X_1, \dots, X_n)$ 的特征函数, 假设 $\vec{X}$ 在 $V = \{a_i \leq X_i \leq b_i, i = 1, \dots, n\}$ 的表面上取值概率为 θ , 则

$$\begin{aligned}&P(a_k \leq X_k \leq b_k, k = 1, 2, \dots, n) \\ &= \lim_{\tau_k \rightarrow +\infty, k=1, \dots, n} \frac{1}{(2\pi)^n} \int_{-\tau_1}^{\tau_1} \dots \int_{-\tau_n}^{\tau_n} \prod_{k=1}^n \frac{e^{-ia_k t_k} - e^{-ib_k t_k}}{it_k} \varphi(t_1, \dots, t_n) dt_1 \dots dt_n\end{aligned}$$

例 3.2.4 $\vec{X} \sim N(\vec{\mu}, \Sigma)$,

$$\begin{aligned}\varphi(t_1, \dots, t_n) &= E[e^{i \sum_{k=1}^n t_k X_k}] \\ Y &= \sum_{k=1}^n t_k X_k \sim N\left(\sum_{k=1}^n t_k \mu_k, \vec{t} \Sigma \vec{t}^T\right) \\ \varphi_r(s) &= E[e^{isY}] = E[e^{is \sum_{k=1}^n t_k X_k}] = \exp\left(is \sum_{k=1}^n t_k \mu_k - \frac{1}{2} s^2 \vec{t} \Sigma \vec{t}^T\right) \\ \varphi(t_1, \dots, t_n) &= \varphi_Y(1) = \exp\left(i \vec{t} \vec{\mu}^T - \frac{1}{2} \vec{t} \Sigma \vec{t}^T\right)\end{aligned}$$

(协方差矩阵相关性质) 设 n 阶协方差矩阵 Σ 的秩为 r ,

1° $n=1, r=0, \sigma^2=0, P(x=\mu)=1$

2° $n=2, r=0, P((X_1, X_2) = (\mu_1, \mu_2))=1$

3° $n=2, r=1, (X_1, X_2)$ 在平面内某条曲线上取值。

$\text{rank} \Sigma < n, (X_1, \dots, X_n)$ 退化到 $\mathbb{R}^n$ 的一个 r 维子空间内取值。

定理 3.2.8 $\varphi(t_1, \dots, t_n)$ 是 $(X_1, \dots, X_n)$ 的特征函数, $E[X_1^{k_1} \dots X_n^{k_n}]$ 存在, 记 $s = k_1 + \dots + k_n$, 则

$$\left. \frac{\partial^s \varphi}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} \right|_{\vec{t}=\vec{0}} = i^s E[X_1^{k_1} \dots X_n^{k_n}]$$

定理 3.2.9 $X_1, \dots, X_n$ 相互独立 $\Leftrightarrow \varphi(t_1, \dots, t_n) = \prod_{k=1}^n \varphi_{X_k}(t_k)$.

3.2.4 函数列的收敛性

定义 3.2.4 如果一系列分布函数 $F_n(x)$ 以及分布函数 $F(x)$, 在 $F(x)$ 的连续点处成立

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

则称 $F_n(x)$ 弱收敛于 $F(x)$, 记作

$$F_n(x) \xrightarrow{w} F(x)$$

分布函数列弱收敛于分布函数, 极限是唯一的。

定理 3.2.10 (连续性定理)

(1). $F_n(x) \xrightarrow{w} F(x)$, 则

$$\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi_t$$

(2). $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi_t$ 且 $\varphi(t)$ 在 $t=0$ 处连续, 则 $\varphi(t)$ 也是某分布函数 $F(x)$ 的特征函数, 且 $F_n(x) \xrightarrow{w} F(x)$.

例 3.2.5 连续性定理 (2) 中 $\varphi(t)$ 在 $t=0$ 处连续的条件不能去掉, 反例:

$$X_n \sim N(0, n), \varphi_n(t) = e^{-\frac{nt^2}{2}} \rightarrow \begin{cases} 1, & t=0 \\ 0, & t \neq 0 \end{cases}$$

后者不是特征函数。

3.3 极限定理简要介绍

定义 3.3.1 (大数定律) 我们知道, 多次独立重复实验中事件 A 的发生频率, 稳定于 $P(A)$

$$\frac{\sum_{j=1}^n X_j}{n} \rightarrow P(A)$$

如果以概率收敛, 称 X_n 服从弱大数定律; 几乎处处收敛, 称 X_n 服从强大数定律。

定义 3.3.2 (依分布收敛) 如果 $F_n(x) \xrightarrow{w} F(x)$, 则称 X_n 依分布收敛于 X , 记作 $X_n \xrightarrow{D} X$.

定理 3.3.1 X_n 独立同分布, $E[X_i] = \mu$, $s_n = X_1 + \cdots + X_n$, 则

$$\frac{s_n}{n} \xrightarrow{D} \mu$$

定理 3.3.2 (中心极限定理) X_n 独立同分布, $E[X_i] = \mu$, $\text{var}(X_i) = \sigma^2$, $s_n = X_1 + \cdots + X_n$, 则

$$\frac{s_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{D} Y, Y \sim N(0, 1)$$

证明 Y 的特征函数为 $\varphi_Y(t) = e^{-\frac{t^2}{2}}$.

令 $Z_k = \frac{X_k - \mu}{\sqrt{n\sigma^2}}$, 其特征函数为 $\varphi_Z(t) = 1 - \frac{t^2}{2n} + o(\frac{t^2}{n})$, 这是因为

$$E[Z_k] = 0, E[Z_k^2] = \frac{1}{n}$$

$\frac{s_n - n\mu}{\sqrt{n\sigma^2}} = \sum_{k=1}^n Z_k$ 的特征函数就是

$$\varphi_n(t) = (\varphi_Z(t))^n = \left(1 - \frac{t^2}{2n} + o(\frac{t^2}{n})\right) \rightarrow e^{-\frac{t^2}{2}}$$

由连续性定理, $\frac{s_n - n\mu}{\sqrt{n\sigma^2}}$ 的分布函数收敛于 $e^{-\frac{t^2}{2}}$ 对应的分布函数。

第四章 极限定理

4.1 收敛性

4.1.1 收敛性的定义

定义 4.1.1

- (1). 几乎处处收敛: 若 $P(\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1$, 记作 $X_n \xrightarrow{\text{a.s.}} X$.
- (2). r 阶平均收敛: $E[|X_n|^r] < +\infty$, 若 $\lim_{n \rightarrow \infty} E[|X_n - X|^r] = 0$. 记作 $X_n \xrightarrow{r} X$.
- (3). 依概率收敛: 若 $\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0$, 记作 $X_n \xrightarrow{P} X$.
- (4). 依分布收敛: 分布函数弱收敛。记作 $X_n \xrightarrow{D} X$.

4.1.2 四种收敛性之间的关系

$$\begin{array}{c} X_n \xrightarrow{\text{a.s.}} X \\ \Downarrow \\ X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X \\ \Uparrow \\ X_n \xrightarrow{r} X \end{array}$$

定理 4.1.1 $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$.

证明 设 X_n 的分布函数为 $F_n(x)$, X 的分布函数为 $F(x)$,

$$\begin{aligned} \forall \varepsilon > 0, F_n(x) &= P(X_n \leq x) = P(X_n \leq X, X \leq x + \varepsilon) + P(X_n \leq x, X > x + \varepsilon) \\ &= P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon) = F(x + \varepsilon) + P(|X_n - X| > \varepsilon) \\ F(x - \varepsilon) &= P(X \leq x - \varepsilon) = P(X \leq x - \varepsilon, X_n \leq x) + P(X \leq x - \varepsilon, X_n > x) \\ &\leq P(X_n \leq X) + P(|X_n - X| > \varepsilon) = F_n(x) + P(|X_n - X| > \varepsilon) \end{aligned}$$

所以 $F(x - \varepsilon) - P(|X_n - X| > \varepsilon) \leq F_n(x) \leq F(x + \varepsilon) + P(|X_n - X| > \varepsilon)$, 令 $n \rightarrow \infty$, 就得到

$$F(x - \varepsilon) \leq \liminf_{n \rightarrow \infty} F_n(x) \leq \limsup_{n \rightarrow \infty} F_n(x) \leq F(x + \varepsilon)$$

若 x 是 $F(x)$ 的连续点, 令 $\varepsilon \rightarrow 0^+$, 就有 $\lim_{n \rightarrow \infty} F_n(x) = F(x)$, 结论得证。

反过来不成立, 例如: ω 为抛硬币正反, 正则 $X(\omega) = 1$, 反则 $X(\omega) = -1$, 令 $X_n = -X$, 此时 X_n 和 X 的分布相同, 但 $|X_n - X| = 2$, 故 $X_n \not\xrightarrow{P} X$

引理 4.1.1 (Markov 不等式) $E[|X|^r] < \infty$, 则对于 $\forall a > 0$,

$$P(|X|^r > a) \leq \frac{E[|X|^r]}{a}$$

证明 由

$$|x|^r \geq a \cdot I_{(|x|^r > a)}$$

得

$$E[|X|^r] \geq E[a \cdot I_{(|X|^r > a)}] = aP(|X|^r > a)$$

定理 4.1.2 $X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{P} X, r \geq 1$.

证明 $\lim_{n \rightarrow \infty} E[|X_n - X|^r] = 0, r \geq 1$

$$\forall \varepsilon, P(|X_n - X| > \varepsilon) = P(|X_n - X|^r > \varepsilon^r) \stackrel{\text{Markov}}{\leq} \frac{E[|X_n - X|^r]}{\varepsilon^r} \rightarrow 0$$

引理 4.1.2 以下结论等价:

1° $X_n \xrightarrow{\text{a.s.}} X$

2°

$$P\left(\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{|X_m - X| > \frac{1}{k}\right\}\right) = 0$$

3°

$$\forall \varepsilon > 0, P\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{|X_m - X| > \varepsilon\}\right) = 0$$

4°

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P\left(\bigcup_{m=n}^{\infty} \{|X_m - X| > \varepsilon\}\right) = 0$$

定理 4.1.3 $X_n \xrightarrow{\text{a.s.}} X \Rightarrow X_n \xrightarrow{P} X$.

证明 $\forall \varepsilon > 0, P(|X_n - X| > \varepsilon) \leq P(\bigcup_{m=n}^{\infty} |X_m - X| > \varepsilon) \rightarrow 0$.

反之不成立, 例如: $\Omega = (0, 1)$,

$$\begin{aligned} X_1 &= I_{(0,1)} \\ X_2 &= I_{(0, \frac{1}{2})}, X_3 = I_{(\frac{1}{2}, 1)} \\ X_4 &= I_{(0, \frac{1}{3})}, X_5 = I_{(\frac{1}{3}, \frac{2}{3})}, X_6 = I_{(\frac{2}{3}, 1)} \\ &\vdots \end{aligned}$$

$P(|X_n - 0| > \varepsilon) = \frac{1}{k} \rightarrow 0$, 故 $X_n \xrightarrow{P} X$, 但 $X_n \not\xrightarrow{\text{a.s.}} X$ 不成立。

定理 4.1.4

(1). $X_n \xrightarrow{D} C$ (常数), 则 $X_n \xrightarrow{P} C$.

(2). $X_n \xrightarrow{P} X$, 且存在常数 k , 使得 $P(|X_n| \leq k) = 1, \forall n$ 则 $X_n \xrightarrow{r} X, r \geq 1$.

(3). 若 $\forall \varepsilon \geq 0$, 都有

$$\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty$$

则 $X_n \xrightarrow{\text{a.s.}} X$.

引理 4.1.3 $X_n \xrightarrow{P} X$ 当且仅当 $\{X_n\}$ 的任意子列 $\{X_{n_m}\}$, 可以找到几乎处处收敛的子列 $\{X_{n(m_k)}\}_{k=1}^{\infty}$.

定理 4.1.5 (Skorokhod 表示定理) $\{X_n\}, X$ 在概率空间 (Ω, F, P) 上的分布函数分别为 $\{F_n(x)\}, F(x)$ 且 $X_n \xrightarrow{D} X$, 则:

存在另一个¹概率空间 (Ω', F', P') 以及在此空间上定义的随机变量 Y_n, Y 满足 Y_n 的分布函数是 $F_n(x)$, Y 的分布函数是 $F(x)$, 且 $Y_n \xrightarrow{\text{a.s.}} Y$.

定理 4.1.6 $X_n \xrightarrow{P} X$ 当且仅当存在有界连续函数 $g(x)$ 满足 $\lim_{n \rightarrow \infty} E[g(X_n)] = E[g(X)]$.

4.2 辅助结论

4.2.1 不等式

(Markov, Chebyshev) $h(x)$ 是非负可测函数, $a > 0$, X 是随机变量, 则

$$P(|x| \geq a) \leq \frac{E[|x|^r]}{a^r} \quad \text{Markov}$$

$$P(|X - E(X)| \geq a) \leq \frac{E[|X - E(X)|^2]}{a^2} = \frac{\text{var}(X)}{a^2} \quad \text{Chebyshev}$$

(Holder) $p, q > 1, \frac{1}{p} + \frac{1}{q} = 1, E(|XY|) \leq (E|X|^p)^{\frac{1}{p}} \cdot (E|Y|^q)^{\frac{1}{q}}$

(Minkovski) $(E|X + Y|^p)^{\frac{1}{p}} \leq (E|X|^p)^{\frac{1}{p}} + (E|Y|^p)^{\frac{1}{p}}$

(no name) $E(|X + Y|^r) \leq C_r(E|X|^r + E|Y|^r)$, 其中 $C_r = \begin{cases} 2^{r-1} & , r > 1 \\ 1 & , 0 < r \leq 1 \end{cases}$

4.2.2 运算性质

定理 4.2.1 $X_n \xrightarrow{*} X, Y_n \xrightarrow{*} Y, * \in \{\text{a.s.}, p, r\}$, 则

(1). 如果 $X_n \xrightarrow{*} Y$, 则 $P(X = Y) = 1$.

(2). $X_n + Y_n \xrightarrow{*} X + Y, X_n Y_n \xrightarrow{*} XY$

定理 4.2.2 如果 $X_n \xrightarrow{D} X, Y_n \xrightarrow{P} C$ (常数), 则 $X_n + Y_n \xrightarrow{D} X + C$.

¹可以与原空间相同。

²后者对于 $* = r$ 的情况不成立。

4.2.3 Borel-Cantelli 引理

定义 4.2.1 对一系列事件 $\{A_n\}$, 定义

$$\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \limsup_{n \rightarrow \infty} A_n = \{A_n \text{ i.o.}\}$$

即发生了无数次事件全体。

$$\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m = \liminf_{n \rightarrow \infty} A_n = \{A_n^c \text{ i.o.}\}^c$$

即只有有限次没发生的事件全体。

定理 4.2.3 (Borel-Cantelli 引理)

- (1). 如果 $\sum_{n=1}^{\infty} P(A_n) < \infty$, 则 $P(A_n \text{ i.o.}) = 0$.
- (2). 若 $\{A_n\}$ 相互独立, $\sum_{n=1}^{\infty} P(A_n) = \infty \Leftrightarrow P(A_n \text{ i.o.}) = 1$.

4.3 大数定律

回忆定义 3.3.1. 本章讨论的是 $\{X_n\}$ 何时满足弱、强大数定律, 即何时

$$\frac{s_n}{n} = \frac{\sum_{k=1}^n X_k}{n}$$

依概率收敛、几乎处处收敛。

4.3.1 弱大数定律

定理 4.3.1 (Khinchin LLN) $\{X_n\}$ 独立同分布, $E[X_n] = \mu$, $s_n = X_1 + \cdots + X_n$, 如果 $\frac{s_n}{n} - \mu = \frac{s_n - E(s_n)}{n} \xrightarrow{D} 0$, 则 $\frac{s_n}{n} - \mu \xrightarrow{P} 0$.

证明 由定理 4.1.4(1) 直接得证。

定理 4.3.2 (Bernoulli LLN) $X_n(\omega) \in \{0, 1\}$, $P(X_n = 1) = p$, $P(X_n = 0) = 1 - p$. 则

$$\frac{s_n}{n} \xrightarrow{P} p(X_n = 1) = p$$

定理 4.3.3 (Chebyshev LLN) $\{X_n\}$ 两两不相关, $\text{var}(X_n) < k$, 则

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

证明

$$\begin{aligned} \forall \varepsilon > 0, P\left(\left|\frac{s_n - E(s_n)}{n}\right| > \varepsilon\right) &\leq \frac{1}{\varepsilon^2} E\left[\left(\frac{s_n - E(s_n)}{n}\right)^2\right] = \frac{1}{n^2 \varepsilon^2} \sum_{i=1}^n E[(X_i - E(X_i))^2] \\ &\leq \frac{k}{n \varepsilon^2} \rightarrow 0 \end{aligned}$$

定理 4.3.4 (Markov LLN) $\frac{1}{n^2} \text{var}(s_n) \rightarrow 0$, 则

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

定理 4.3.5 X_n 独立, 设

$$Y_{n,k} = \begin{cases} X_k & , |X_k| \leq n \\ 0 & , |X_k| > n \end{cases}$$

记 $a_n = \sum_{k=1}^n E(Y_{n,k})$, $b_n = n$, 则:

$$\sum_{k=1}^n P(|X_k| > n) \rightarrow 0$$

$$\frac{1}{n^2} \sum_{k=1}^n E(Y_{n,k}^2) \rightarrow 0 \Rightarrow \frac{s_n - a_n}{b_n} \xrightarrow{P} 0$$

4.3.2 强大数定律

定理 4.3.6 $\{X_n\}$ 独立同分布, $E(X_1^2) < \infty$, $E[X_i] = \mu$, 则

- (1). $\frac{s_n}{n} \xrightarrow{2} \mu$.
- (2). $\frac{s_n}{n} \xrightarrow{\text{a.s.}} \mu$.

定理 4.3.7 (Kolmogorov) $\{X_n\}$ 相互独立, 且

$$\sum_{n=1}^{\infty} \frac{\text{var}(X_n)}{n^2} < \infty$$

则

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

定理 4.3.8 $\{X_n\}$ 独立同分布,

$$\frac{s_n}{n} - \mu \xrightarrow{\text{a.s.}} 0 \Leftrightarrow E[X_n] = \mu$$

4.4 中心极限定理

在第 3.3 节曾介绍过中心极限定理, 其中要求了 X_n 独立同分布。实际上, 这个条件还可以减弱。

(Feller 条件) $\sigma_k^2 = \text{var}(X_k)$, 满足

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{\sigma_1^2 + \cdots + \sigma_n^2} = 0$$

(Lindeberg 条件) $\{X_n\}$ 独立, 记 $a_k = E[X_k]$, $b_k^2 = \text{var}(X_k)$, $B_n^2 = \sum_{k=1}^n b_k^2$, $F_k(x)$ 是 X_k 的分布函数, 满足: 对于任意 $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-a_k| > \varepsilon B_n} (x-a_k)^2 dF_k(x) = 0$$

等价于

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n E((X_k - a_k)^2 I_{\{|X_k - a_k| > \varepsilon B_n\}}) = 0$$

引理 4.4.1 Feller $\Rightarrow \max_{1 \leq k \leq n} \left| \frac{x_k - a_k}{B_n} \right| \xrightarrow{P} 0$.

引理 4.4.2 Lindeberg $\Rightarrow$ Feller.

定理 4.4.1 (Lindeberg-Feller CLT) $\{X_n\}$ 独立, 满足 Lindeberg 条件, 则

$$\frac{s_n - E[s_n]}{B_n} \xrightarrow{D} N(0, 1)$$

定理 4.4.2 (Lyapunov CLT) $\{X_n\}$ 独立, 若存在 $\delta > 0$ s.t.

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

则

$$\frac{s_n - E[s_n]}{B_n} \xrightarrow{D} N(0, 1)$$