



初等概率论笔记

2024-2025 学年秋冬学期

作者: Yishu Jiang

组织: School of Economics, Zhejiang University

时间: September 26, 2024



不要以为抹消过去, 重新来过, 即可发生什么改变。——比企谷八幡

前言

本笔记是 2024-2025 学年秋冬学期 3.5 学分概率论的学习笔记, 基于张立新老师上课的讲义整理而成, 同时参考了何书元老师的《概率论》、李贤平老师的《概率论基础》、吴昊老师的《概率论 (I)》讲义、中科大概率论习题课讲义. 目前的基本架构是分成三部分: 第一部分侧重于概念的介绍和对为什么引入的解释; 第二部分侧重于补充内容的整理; 第三部分是习题集. 后续如果还有机会的话, 会再加入测度论视角的概率论内容.

由于本人初次学习概率论且水平能力有限, 笔记中有所疏漏处, 恳请指正.

目录

第一部分 初等概率论	1
第 1 章 概率空间	2
1.1 概率的公理化定义	2
1.2 概率的连续性	2
1.3 Borel-Canteli 引理	3
1.4 条件概率、全概率公式、Bayes 公式	4
1.4.1 条件概率	4
1.4.2 全概率公式与 Bayes 公式	4
第 2 章 随机变量	5
2.1 随机变量的定义	5
2.2 离散型随机变量	5
2.3 连续型随机变量	5
2.4 常见随机变量的分布	5
第 3 章 随机向量	6
3.1 随机向量	6
3.2 离散型随机向量	7
3.3 连续型随机向量	8
第 4 章 重要数据特征	10
4.1 数学期望与方差	10
4.2 数学期望与方差的性质	11
4.3 协方差、相关系数	13
4.4 条件数学期望	13
第 5 章 收敛性与中心极限定理	14
5.1 矩母函数	14
5.2 特征函数	14
5.3 多元正态分布	14
5.4 大数定律	14
5.5 中心极限定理	14
第二部分 补充内容	15
第三部分 习题整理 (含小测题、历年题)	16
第 6 章 茆书上习题	17
第 7 章 吴昊讲义上习题	18

第 8 章 小测题	19
8.1 第一次小测	19
8.2 第二次小测	20
8.3 第三次小测	21
第 9 章 历年卷	22
9.1 2023-2024 概率论 3.5 学分 zlx 班	22
9.2 2022-2023 概率论 3.5 学分 zlx 班	23
9.3 2021-2022 概率论 3.5 学分 zlx 班	24

第一部分

初等概率论

第 1 章 概率空间

在讨论概率空间之前, 首先我们需要用对集合的运算刻画事件——在实际情况中往往因为信息问题, 我们只能知道其中一部分事件的行为, 而需要刻画所有事件的概率, 我们首先需要找到一个能对事件的运算封闭的集合.

1.1 概率的公理化定义

1.2 概率的连续性

类似函数极限 $\lim_{n \rightarrow \infty} f(x_n) = f(\lim_{n \rightarrow \infty} x_n)$, 我们不禁要问, 概率测度是否也具有这样良好的性质, 即允许极限与概率测度换序.

从简单的事件列研究起, 由于概率空间中的事件一定是“有界”的 ($\emptyset \subset A \subset \Omega$), 参照数列极限, 我们首先考虑单调序列: 对于单调增的事件列 $\{A_n\}_{n=1}^{\infty}$, 有 $A_1 \subset A_2 \subset \dots$, 则有 $\bigcup_{i=1}^{\infty} A_i = \lim_{n \rightarrow \infty} A_n$; 同理, 对于单调减的事件列 $\{B_n\}_{n=1}^{\infty}$, 指 $B_1 \supset B_2 \supset \dots$, 则有 $\bigcap_{i=1}^{\infty} B_i = \lim_{n \rightarrow \infty} B_n$. 基于此, 我们可以得到:

定理 1.1 (概率的上、下连续性)

对于单调增的事件列 $\{A_n\}_{n=1}^{\infty}$, 有 $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(\lim_{n \rightarrow \infty} A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} A_n)$, 称上连续性.

对于单调减的事件列 $\{B_n\}_{n=1}^{\infty}$, 有 $\lim_{n \rightarrow \infty} \mathbb{P}(B_n) = \mathbb{P}(\lim_{n \rightarrow \infty} B_n) = \mathbb{P}(\bigcap_{n=1}^{\infty} B_n)$, 称下连续性.

证明 首先考虑单调增的事件列的情形. 我们希望能使用概率的“可列可加性”, 因此需要把 $\{A_n\}$ 转化成一列相互不相交的事件列. 做如下变换:

$$B_k = A_k - A_{k-1}, k = 1, 2, \dots; A_0 = \emptyset$$

则 $\{B_k\}$ 相互不相交且有 $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ (类似于在 Venn 图里拆成“一圈一圈的饼”).

故 $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) = \sum_{n=1}^{\infty} (\mathbb{P}(A_n) - \mathbb{P}(A_{n-1})) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n)$.

而对于单调减的事件列, 只需要取单调增情形时的余集, 即可套用上面的过程得到结论.

现在考虑一般的事件列 $\{A_n\}$, 在没有单调性的情形下, 再借鉴数列极限中的上下极限的概念, 我们定义事件列的上下极限.

定义 1.1 (事件列的上下极限)

对于事件列 A_1, A_2, \dots 设

$$\begin{cases} C_n = \bigcup_{j=n}^{\infty} A_j \\ D_n = \bigcap_{j=n}^{\infty} A_j \end{cases} \quad \text{易知 } C_n \text{ 单调递减, } D_n \text{ 单调递增, 则极限存在.}$$
$$\lim_{n \rightarrow \infty} C_n = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j =: \overline{\lim}_{n \rightarrow \infty} A_n (\sup_{n \geq 1} \inf_{m \geq n} A_m)$$
$$\lim_{n \rightarrow \infty} D_n = \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} A_j =: \underline{\lim}_{n \rightarrow \infty} A_n (\inf_{n \geq 1} \sup_{m \geq n} A_m)$$

从定义中, 可以看出: 下极限是上极限的子集, 即 $\underline{\lim}_{n \rightarrow \infty} A_n \subset \overline{\lim}_{n \rightarrow \infty} A_n$. 于是可以定义: $\lim_{n \rightarrow \infty} A_n$ 存在当且仅当 $\underline{\lim}_{n \rightarrow \infty} A_n = \overline{\lim}_{n \rightarrow \infty} A_n$

更进一步地理解, 可以发现 $\overline{\lim}_{n \rightarrow \infty} A_n$ 发生当且仅当有无穷个 A_j 发生, 故而也被记作 $\{A_n \text{ i.o.}\}$ (infinitely often); $\underline{\lim}_{n \rightarrow \infty} A_n$ 发生当且仅当最多有限个 A_j 不发生.

在此基础上, 我们证明概率测度的连续性: $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(\lim_{n \rightarrow \infty} A_n)$.

引理 1.1 (Fatou 引理)

$$\begin{aligned}\mathbb{P}(\liminf_{n \rightarrow \infty} A_n) &= \mathbb{P}(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m) = \lim_{n \rightarrow \infty} \mathbb{P}(\bigcap_{m=n}^{\infty} A_m) \leq \lim_{n \rightarrow \infty} \inf \mathbb{P}(A_n) \\ \mathbb{P}(\limsup_{n \rightarrow \infty} A_n) &= \mathbb{P}(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m) = \lim_{n \rightarrow \infty} \mathbb{P}(\bigcup_{m=n}^{\infty} A_m) \geq \lim_{n \rightarrow \infty} \sup \mathbb{P}(A_n)\end{aligned}$$



要使 $\lim_{n \rightarrow \infty} A_n$ 存在, 则有 $\lim_{n \rightarrow \infty} \inf A_n = \lim_{n \rightarrow \infty} \sup A_n = \lim_{n \rightarrow \infty} A_n$. 由夹逼定理知 $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(\lim_{n \rightarrow \infty} A_n)$ 即证明了概率测度的连续性.

更进一步地, 可以证明, 概率测度的“可列可加性”等价于“有限可加性 + 连续性”.

命题 1.1

设 $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$ 满足有限可加性, 且对 $\{A_n\} \subset \mathcal{F}, A_n \searrow \emptyset$ 有 $\mathbb{P}(A_n) \rightarrow 0$. 则 \mathbb{P} 满足可列可加性.



证明 考虑事件列 $\{B_n\}_{n=1}^{\infty}$ 满足 $B_i \cap B_j = \emptyset, \forall i, j$ 且 $B_n \searrow$.

令 $B = \sum_{i=1}^{\infty} B_i = \sum_{i=1}^n B_i + \sum_{i=n+1}^{\infty} B_i$, 并设 $C_n = \sum_{i=n+1}^{\infty} B_i$ 则 $C_n \searrow \emptyset$.

于是 $\mathbb{P}(B) = \mathbb{P}(\sum_{i=1}^n B_i) + \mathbb{P}(C_n) \rightarrow \mathbb{P}(\sum_{i=1}^n B_i)$ as $n \rightarrow \infty$.

1.3 Borel-Canteli 引理

重新回顾一下集合列的上下极限.

定义 1.2 (事件列的上下极限)

对于事件列 A_1, A_2, \dots 设

$$\begin{cases} C_n = \bigcup_{j=n}^{\infty} A_j \\ D_n = \bigcap_{j=n}^{\infty} A_j \end{cases} \quad \text{易知 } C_n \text{ 单调递减, } D_n \text{ 单调递增, 则极限存在.}$$

$$\lim_{n \rightarrow \infty} C_n = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j =: \overline{\lim}_{n \rightarrow \infty} A_n$$

$$\lim_{n \rightarrow \infty} D_n = \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} A_j =: \underline{\lim}_{n \rightarrow \infty} A_n$$



在定义了事件列上下极限的基础上, 引入 Borel-Canteli 引理.

定理 1.2 (Borel-Canteli 引理)

对于事件列 A_1, A_2, \dots

(1): 若 $\sum_{j=1}^{\infty} \mathbb{P}(A_j) < \infty$, 则 $\mathbb{P}(A_n \text{ i.o.}) = 0$;

(2): 若事件列中事件相互独立, $\sum_{j=1}^{\infty} \mathbb{P}(A_j) = \infty$, 则 $\mathbb{P}(A_n \text{ i.o.}) = 1$.



思路:

(1). $\mathbb{P}(A_n \text{ i.o.}) = \lim_{n \rightarrow \infty} \mathbb{P}(\bigcup_{j=n}^{\infty} A_j)$, 考虑到条件给的是事件列概率的和, 故而我们可以想到通过次可列可加性, 将事件列并的概率放缩成概率列的部分和, 再通过正项级数收敛的性质进行估计.

(2). $\mathbb{P}(A_n \text{ i.o.}) = \lim_{n \rightarrow \infty} \mathbb{P}(\bigcup_{j=n}^{\infty} A_j)$, 对 $\mathbb{P}(\bigcup_{j=n}^{\infty} A_j)$ 进行估计, 由于要使用到事件相互独立的条件, 我们可以通过 De-Morgan 定律将并改写为交, 即得到 $\mathbb{P}(\bigcup_{j=n}^{\infty} A_j) = \mathbb{P}(\overline{\bigcap_{j=n}^{\infty} \overline{A_j}}) = 1 - \mathbb{P}(\bigcap_{j=n}^{\infty} \overline{A_j}) = 1 - \prod_{j=n}^{\infty} \mathbb{P}(\overline{A_j}) = 1 - \prod_{j=n}^{\infty} (1 - \mathbb{P}(A_j))$. 到此为止, 我们用完了事件独立性的条件, 接下来的目标就是再进行估计, 把概率和的条件用上, 就需要把乘积转换成和的形式, 考虑如下放缩 $1 - x \leq e^{-x}, x \in [0, 1]$, 于是上式可以改写成指数上带求和的形式, 从而得到收敛.

证明 (1): 由 $\mathbb{P}(A_n \text{ i.o.}) = \lim_{n \rightarrow \infty} \mathbb{P}(\bigcup_{j=n}^{\infty} A_j)$, 又 $\mathbb{P}(\bigcup_{j=n}^{\infty} A_j) \leq \sum_{j=n}^{\infty} \mathbb{P}(A_j)$, 得 $\mathbb{P}(A_n \text{ i.o.}) \leq \lim_{n \rightarrow \infty} \sum_{j=n}^{\infty} \mathbb{P}(A_j)$. 而 $\sum_{j=1}^{\infty} \mathbb{P}(A_j) < \infty$, 则根据级数收敛的 Cauchy 准则知 $\lim_{j \rightarrow \infty} \sum_{j=n}^{\infty} \mathbb{P}(A_j) = 0$. 故 $\mathbb{P}(A_n \text{ i.o.}) = 0$.

(2): 由 $\mathbb{P}(A_n \text{ i.o.}) = \lim_{n \rightarrow \infty} \mathbb{P}(\bigcup_{j=n}^{\infty} A_j)$, 又

$$\mathbb{P}(\bigcup_{j=n}^{\infty} A_j) = \mathbb{P}(\overline{\bigcap_{j=n}^{\infty} \overline{A_j}}) = 1 - \mathbb{P}(\bigcap_{j=n}^{\infty} \overline{A_j}) = 1 - \prod_{j=n}^{\infty} \mathbb{P}(\overline{A_j}) = 1 - \prod_{j=n}^{\infty} (1 - \mathbb{P}(A_j))$$

考虑 $1 - x \leq e^{-x}$, $0 \leq x \leq 1$, 有 $\mathbb{P}(\bigcup_{j=n}^{\infty} A_j) \leq 1 - \prod_{j=n}^{\infty} e^{-\mathbb{P}(A_j)} = 1 - e^{-\sum_{j=n}^{\infty} \mathbb{P}(A_j)}$. 由于 $\sum_{j=1}^{\infty} \mathbb{P}(A_j) = \infty$, 则 $-\sum_{j=n}^{\infty} \mathbb{P}(A_j) = -\infty$, 故可以得到 $\mathbb{P}(\bigcup_{j=n}^{\infty} A_j) = 1$, 于是有 $\mathbb{P}(A_n \text{ i.o.}) = 1$.

注:

1. 在事件列中事件是相互独立时, $\mathbb{P}(A_n \text{ i.o.})$ 只能取 0 或 1, 即对相互独立的事件列, 其中有无穷个事件发生的概率要么为 0 要么为 1. 这一结论被称为独立 0-1 律.
2. 对于独立重复试验 (每次发生的概率为 p), 有 $\sum_{n=1}^{\infty} \mathbb{P}(A_n) \leq \infty$ 当且仅当 $p = 0$. 也就是说, 在独立重复试验中, 当事件几乎必然不发生时, 无穷次地重复下去几乎必然不会有无穷次事件发生; 当事件每次发生的概率不为 0 时, 无论这个概率多么地小, 只要无穷次地重复下去几乎必然会有无穷次事件发生.
3. 对于独立试验 (第 n 次发生的概率为 $p_n (> 0)$), 无穷次地做下去, 事件列中是否有无穷多个事件发生只和 n 充分大时 p_n 的值有关 (这是符合直觉的). 然而, 直觉所感受不到的是: $p_n = \frac{1}{n}$ 时, 事件列中会有无穷多个事件发生; $p_n = \frac{1}{n^2}$ 时, 则不会有无穷多个事件发生.
4. Borel-Canteli 引理是强大数定律的基础 (参见之后的强大数定律).

1.4 条件概率、全概率公式、Bayes 公式

1.4.1 条件概率

条件概率的本质就是对样本空间的限制.

定义 1.3 (条件概率)



1.4.2 全概率公式与 Bayes 公式

定理 1.3 (全概率公式)

事件 $\{A_i\}_{i=1}^n$ 互斥, $B \subset \bigcup_{i=1}^n A_i$, 则

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B | A_i) \mathbb{P}(A_i)$$



证明很简单, $B = \bigcup_{i=1}^n B A_i$ 再利用概率的可列可加性并用条件概率的乘法公式处理即可.

以下是两个常见的应用.

例题 1.1 (抽签公平性) n 个球, 有 m 个黑球, 剩下全为白, 球除了颜色外没有任何差别. 求证: 无放回地依次抽取球, 每一次抽中黑球的概率都是 $\frac{m}{n}$.

再来一个更加有趣的: 赌徒破产模型, 同时在这里也会介绍一种在概率论中非常常见的处理方法: 递推公式法.

例题 1.2 (赌徒破产模型) 一个人有 a 的本金, 打算再赢 b 元就停止赌博, 设每局 $p = \frac{1}{2}$ 概率赢, 输赢对金钱的影响都是 1, 输光后自然地停止赌博, 求输光的概率 $p(a)$.

第 2 章 随机变量

2.1 随机变量的定义

2.2 离散型随机变量

2.3 连续型随机变量

2.4 常见随机变量的分布

第3章 随机向量

本章主要涉及随机向量及其分布,从本质上看,随机向量与随机变量差异不大,只是在处理方式上更加复杂了.

3.1 随机向量

定义 3.1 (随机向量)

已知概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$, X_1, \dots, X_n 为此概率空间上的随机变量, 则 $\vec{X} = (X_1, \dots, X_n)$ 称为 n 维随机向量.



事实上, 随机向量就是对随机变量做了升维处理, 由此我们不仅可以处理 X, Y 各自的分布, 还可以研究这两个分量之间的联系. 在本节中, 我们的主要目标是将在上一章中随机变量的性质和一些处理方法推广到随机向量上; 同时, 由于随机向量的结构更加复杂, 我们还能得到一些在随机变量中得不到的结论.

在对随机变量的讨论中, 我们了解到了分布函数的重要性: 它与概率分布相互唯一决定. 于是我们也同样从分布函数入手, 研究随机向量, 类似的, 定义随机向量的分布函数如下:

定义 3.2 (随机向量的联合分布函数)

$\vec{X} = (X_1, \dots, X_n)$ 为随机向量, 则

$$F(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$$



为了简化处理, 我们默认研究的是二维的随机向量 (X, Y) , 更高维的情形类似. 很自然地推广一维情形下的分布函数的性质:

1. $0 \leq F(x, y) \leq 1$
2. F 分别关于 x, y 为单调递增的右连续函数
3. $F(-\infty, y) = 0, F(x, -\infty) = 0, F(-\infty, -\infty) = 0; F(+\infty, +\infty) = 1$

然而, 联合分布描述的是多个随机变量混合在一块儿发挥作用时的分布; 我们不禁要问: 给定联合分布函数, 能否将其中一个或多个随机变量分离出来, 单独考察他们的随机分布呢?

考虑联合分布函数 $F(x, y) = \mathbb{P}(X \leq x, Y \leq y)$, 如果要求 X 的概率分布, 那就需要让 Y 在这一联合分布中失去作用, 换言之, 就是让 $Y \leq y$ 在这儿“不起作用”, 可以令 $y \rightarrow +\infty$ (相当于对 Y 不做限制 $\{Y \in \mathbb{R}\}$), 根据概率的连续性:

$$\lim_{y \rightarrow +\infty} \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x, \bigcup_{y=0}^{+\infty} \{Y \leq y\}) = \mathbb{P}(X \leq x)$$

从而我们引出了边缘分布的定义.

定义 3.3 (随机向量关于某个分量的边缘分布)

随机向量 $\vec{X} = (X_1, \dots, X_n), 1 \leq k \leq n-1, F$ 为联合分布, 则 (X_1, \dots, X_k) 的分布函数

$$\lim_{x_{k+1}, \dots, x_n \rightarrow +\infty} F(x_1, \dots, x_k, x_{k+1}, \dots, x_n)$$

称为 \vec{X} 关于 (X_1, \dots, X_k) 的边缘分布.



有了边缘分布的概念之后, 我们便可以得到关于随机向量内部各个分量之间独立性的定理, 它叙述的是随机向量内部的关系.

定理 3.1 (随机向量分量的独立性)

随机向量 (X, Y) , X, Y 相互独立的充要条件为联合分布等于关于每个分量的边缘分布的乘积, 即

$$F(x, y) = F_X(x)F_Y(y)$$

其中 F 为联合分布, $F_X(x)$ 表示关于 X 的边缘分布.



这个定理的证明很简单, 用独立性的定义即可. 之后还会有很多判定独立性的定理, 但联合分布与边缘分布的关系总是与独立性最本质相关的.

3.2 离散型随机向量

定义 3.4 (离散型随机变量)

当 X_1, \dots, X_n 均为离散型随机变量时, $\vec{X} = (X_1, \dots, X_n)$ 为离散型随机变量.



我们主要关注离散型随机向量的分布列及边缘分布. 为了简化, 这里讨论的都是二维随机向量 (X, Y) . 设 $X = x_1, x_2, \dots; Y = y_1, y_2, \dots$ 为这两个随机变量的所有取值, 且 $\mathbb{P}(X = x_i, Y = y_j) = p_{ij}$. 则有以下简单性质:

1. $0 \leq p_{ij} \leq 1$;
2. $\sum_{i,j} p_{ij} = 1$;
3. $\mathbb{P}(X = x_i) = \sum_j p_{ij}, \mathbb{P}(Y = y_j) = \sum_i p_{ij}$.

这里介绍一个常见的离散型随机向量分布: 多项分布.

定义 3.5 (多项分布)

A_1, A_2, \dots, A_r 为完备事件组. 现在独立重复试验 n 次, X_i 表示 A_i 发生的次数, $\mathbb{P}(A_i) = p_i (i = 1, 2, \dots, r)$, 则 $\mathbb{P}(X_1 = k_1, \dots, X_r = k_r) = \binom{n}{k_1 \dots k_r} p_1^{k_1} \dots p_r^{k_r}$, 其中 $k_i \geq 0, k_1 + k_2 + \dots + k_r = n$. $\binom{n}{k_1 \dots k_r} = \frac{n!}{k_1! k_2! \dots k_r!}$.



最后给出一个有关离散型随机向量分量独立性的判据, 注意只能用在离散型随机向量上:

定理 3.2 (离散型随机向量分量独立性)

(X, Y) 为离散型随机向量, 则 X 与 Y 独立的充要条件为

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j), \forall i, j$$



3.3 连续型随机向量

定义 3.6 (连续型随机向量)

对于随机向量 $\vec{X} = (X_1, \dots, X_n)$, 若存在 \mathbb{R}^n 上的非负可积函数 f , 满足对 $\forall D = \{(x_1, \dots, x_n) : a_i \leq x_i \leq b_i, a_i, b_i \in \mathbb{R}\}$, 都有

$$\mathbb{P}(\vec{X} \in D) = \int_D f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

则称 \vec{X} 为连续型随机向量, f 为联合密度.



很自然的, 有一个问题: 问什么不像离散型随机向量一样, 用每个分量均为连续的随机变量来定义连续型随机向量呢? 先暂时保留这个问题, 在之后我们会看到这样定义是错误的. 在此之前, 先来关注一下随机向量联合密度函数的一些性质, 同样推广一维的情形, 可以得到:

1. $\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \cdots dx_n = 1$;
2. 联合密度函数不唯一.(修改一个零测集上的取值不影响 n 重积分的结果)
3. 若 f, g 都是 \vec{X} 的联合密度, 且均在 $\vec{x} = (x_1, \dots, x_n)$ 处连续, 则必有 $f(\vec{x}) = g(\vec{x})$

关于连续型随机向量的边缘密度, 事实上我们给出边缘分布函数定义时给出的推导思路大同小异.

这里不加证明的给出需要用到的 Fubini 定理.

定理 3.3 (Fubini 定理)

若 $f(x_1, x_2, \dots, x_n)$ 为非负函数或者在 $D \subset \mathbb{R}^n$ 上绝对可积的函数, 则 f 在 D 上的 n 重积分可以任意交换 n 次积分的次序.



联系之前对边缘分布的定义, 采取同样的方法.

$\vec{X} = (X_1, \dots, X_n)$ 为连续型随机向量, f 为联合密度, $\forall 1 \leq k \leq n-1, D_k = \{(x_1, \dots, x_k) : a_i < x_i < b_i, i = 1, 2, \dots, k\}$, 则 $\mathbb{P}((X_1, \dots, X_k) \in D_k) = \int_{D_k \times \mathbb{R}^{n-k}} f(x_1, \dots, x_n) dx_1 \cdots dx_n$. 由 Fubini 定理, 原式 = $\int_{D_k} (\int_{\mathbb{R}^{n-k}} f(x_1, \dots, x_k, x_{k+1}, \dots, x_n) dx_{k+1} \cdots dx_n) dx_1 \cdots dx_k$

定义 3.7 (随机向量的边缘分布函数)

随机向量 $\vec{X} = (X_1, \dots, X_n)$ 关于分量 (X_1, \dots, X_k) 的边缘密度为

$$\int_{\mathbb{R}^{n-k}} f(x_1, \dots, x_k, x_{k+1}, \dots, x_n) dx_{k+1} \cdots dx_n$$



特别解释一下为何密度函数叫“密度”? 可以这样理解: 某点处的密度函数的值代表了这个点附近单位体积中取值的概率.

设 f 为 $\vec{X} = (X_1, \dots, X_n)$ 的联合密度, 且在 $\vec{x}_0 = (x_1^0, \dots, x_n^0)$ 连续.

取充分小的 $(\Delta x_1, \dots, \Delta x_n)$, 考虑 \vec{X} 在 \vec{x}_0 附近的单位体积中的概率

$$\frac{\mathbb{P}(x_1^0 - \frac{\Delta x_1}{2} < X_1 \leq x_1^0 + \frac{\Delta x_1}{2}, \dots, x_n^0 - \frac{\Delta x_n}{2} < X_n \leq x_n^0 + \frac{\Delta x_n}{2})}{\Delta x_1 \Delta x_2 \cdots \Delta x_n} = \frac{\int_{x_1^0 - \frac{\Delta x_1}{2}}^{x_1^0 + \frac{\Delta x_1}{2}} \cdots \int_{x_n^0 - \frac{\Delta x_n}{2}}^{x_n^0 + \frac{\Delta x_n}{2}} f(x_1, \dots, x_n) dx_1 \cdots dx_n}{\Delta x_1 \Delta x_2 \cdots \Delta x_n}$$

由积分中值定理, $\exists \xi_i \in [x_i^0 - \frac{\Delta x_i}{2}, x_i^0 + \frac{\Delta x_i}{2}]$, $i = 1, 2, \dots, n$, 使得

$$\int_{x_1^0 - \frac{\Delta x_1}{2}}^{x_1^0 + \frac{\Delta x_1}{2}} \cdots \int_{x_n^0 - \frac{\Delta x_n}{2}}^{x_n^0 + \frac{\Delta x_n}{2}} f(x_1, \dots, x_n) dx_1 \cdots dx_n = f(\xi_1, \dots, \xi_n) \Delta x_1 \cdots \Delta x_n$$

故 $\frac{\mathbb{P}(x_1^0 - \frac{\Delta x_1}{2} < X_1 \leq x_1^0 + \frac{\Delta x_1}{2}, \dots, x_n^0 - \frac{\Delta x_n}{2} < X_n \leq x_n^0 + \frac{\Delta x_n}{2})}{\Delta x_1 \Delta x_2 \cdots \Delta x_n} = f(\xi_1, \dots, \xi_n)$.

再根据 f 在 \vec{x}_0 处的连续性, 易知 $f(\xi_1, \dots, \xi_n) \rightarrow f(\vec{x}_0)$.

再讨论联合分布与联合密度之间的关系. 同样是对一维情形的推广, 但需要注意新增的细节. 出于简化, 讨论

二维随机向量 (X, Y) .

特别注意: 这里多出了约束条件 $\int_{\mathbb{R}^2} f(x, y) dx dy = 1$. 接下来用反例说明这是必要的.

例题 3.1

定理 3.4 (连续型随机向量的分量间的独立性)

设 X 有边缘密度函数 $f_X(x)$, Y 有边缘密度函数 $f_Y(y)$, 则 X, Y 独立的充要条件是 $f_X(x)f_Y(y)$ 为 (X, Y) 的联合密度.



第4章 重要数据特征

本章主要关注一些重要的数据特征,比如期望、方差、协方差等,包括他们的计算、性质.

引入数学期望的一个动机是,已知概率分布,我们需要一个量来反映随机变量的平均取值.很自然地,如果一个随机变量取某一个值的概率更大,那么这个取值在平均量中产生的作用更大,于是,用概率去做加权平均不失为一种好办法.

在此基础上,我们再引入方差,用以度量随机变量的分布对平均值的靠近程度,即其在平均值周围分布的稀疏程度.很自然的,我们会考虑 $|X - \mu|$ 也就是“距离”去和概率加权,但这样就会有一个问题:绝对值函数存在不可微的情形,而二次函数是处处可微又保持非负的,故而我们用 $(X - \mu)^2$ 去和概率加权平均来定义方差.

4.1 数学期望与方差

定义 4.1 (离散型随机变量的期望)

若 X 为离散型随机变量,概率分布为 $\mathbb{P}(X = x_j) = p_j, j = 1, 2, \dots$. 如果

$$\sum_j |x_j| p_j$$

收敛,则称数学期望存在,且数学期望 $\mathbb{E}X = \sum_j x_j p_j$.



定义 4.2 (连续型随机变量的期望)

若 X 为连续型随机变量,密度为 $f(x)$. 如果

$$\int_{-\infty}^{+\infty} |x| f(x) dx$$

收敛,则称数学期望存在,且数学期望 $\mathbb{E}X = \int_{-\infty}^{+\infty} x f(x) dx$.



需要特别注意的是,这里要用到级数/积分的绝对收敛,这是为了保证,任意两项交换次序,并不会影响数学期望的结果,从而保证如果数学期望存在,则一定是唯一的.

以下是一些常见分布的数学期望.

1. $X \sim B(1, p), \mathbb{E}X = p$
2. $X \sim B(n, p), \mathbb{E}X = np$
3. $X \sim P(\lambda), \mathbb{E}X = \lambda$
4. $X \sim G(p), \mathbb{E}X = \frac{1}{p}$
5. $X \sim Pas(r, p), \mathbb{E}X = \frac{r}{p}$
6. $X \sim H(n, M, N), \mathbb{E}X = n \frac{M}{N}$
7. $X \sim U(a, b), \mathbb{E}X = \frac{a+b}{2}$
8. $X \sim \epsilon(\lambda), \mathbb{E}X = \frac{1}{\lambda}$
9. $X \sim N(\mu, \sigma^2), \mathbb{E}X = \mu$
10. $X \sim \Gamma(\alpha, \beta), \mathbb{E}X = \frac{\alpha}{\beta}$

具体计算的过程如下:

1. 显然
- 2.

定义 4.3 (随机变量的方差与标准差)

若随机变量 X 的数学期望 μ 存在且有限, 则 $Var X = \mathbb{E}(X - \mu)^2$ 称为 X 的方差, 也可以记为 $\mathbb{D}X$ 或者 σ_{XX} .

若方差有限, 则 $\sigma_X = \sqrt{Var X}$ 称为 X 的标准差.



利用数学期望的线性性质, 可以得到

$$\mathbb{E}(X - \mu)^2 = \mathbb{E}(X^2 - 2\mu X + \mu^2) = \mathbb{E}X^2 - 2\mu\mathbb{E}X + \mu^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$$

这是随机变量的期望与方差之间的关系 (可以运用到计算当中, 但因为还要另求 $Y = X^2$ 的期望, 其实很繁琐, 在下一节当中会介绍更加简洁有效的方法).

以下是一些常见分布的方差

1. $X \sim B(1, p), Var X = p(1 - p)$
2. $X \sim B(n, p), Var X = np(1 - p)$
3. $X \sim P(\lambda), Var X = \lambda$
4. $X \sim G(p), Var X = \frac{1-p}{p^2}$
5. $X \sim Pas(r, p), Var X = r \frac{1-p}{p^2}$
6. $X \sim H(n, M, N), Var X = n \frac{M}{N} (1 - \frac{M}{N}) \frac{N-n}{N-1}$
7. $X \sim U(a, b), Var X = \frac{(b-a)^2}{12}$
8. $X \sim \epsilon(\lambda), Var X = \frac{1}{\lambda^2}$
9. $X \sim N(\mu, \sigma^2), Var X = \sigma^2$
10. $X \sim \Gamma(\alpha, \beta), Var X = \frac{\alpha}{\beta^2}$

4.2 数学期望与方差的性质

首先给出一种更加便捷的数学期望的计算方法, 不加证明地给出期望公式.

定理 4.1 (期望公式)

(X, Y) 为随机向量, $g: \mathbb{R}^2 \rightarrow \mathbb{R}$, 有 $Z = g(X, Y)$.

若随机向量为离散型的, 有分布列 $\mathbb{P}(X = x_i, Y = y_j) = p_{ij}$, 则

$$\mathbb{E}Z = \mathbb{E}g(X, Y) = \sum_{i,j} p_{ij} g(x_i, y_j)$$

若随机向量为连续型的, 有联合密度 $f(x, y)$, 则

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



简单来说, 就是将随机变量的函数看成一个整体对原先的概率做加权平均即可. 于是, 如果 X 是连续型随机变量, 有密度函数 $f(x)$, 则要求 $\mathbb{E}X^2$ 即求

$$\int_{-\infty}^{+\infty} x^2 f(x) dx$$

特别的, 这里并不要求 $g(x, y)$ 要显式地含有 x, y . 例如已知 (X, Y) 的联合密度函数为 $f(x, y)$, 现在要求 $\mathbb{E}X$, 只需要求

$$\int \int_{\mathbb{R}^2} x f(x, y) dx dy$$

这比先求边缘分布 $f_X(x)$ 再求 $\mathbb{E}X$ 要方便地多.

例题 4.1 X, Y 独立, 且均服从标准正态分布, $Z = (X^2 + Y^2)^\alpha, \alpha > -1$, 求 $\mathbb{E}Z$.

由连续型随机向量各分量独立的充要条件, $f(x, y) = f_X(x)f_Y(y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$ 为联合密度.

由期望公式, $\mathbb{E}Z = \int \int_{\mathbb{R}^2} f(x, y)(x^2 + y^2)^\alpha dx dy = \frac{1}{2\pi} \int \int_{\mathbb{R}^2} (x^2 + y^2)^\alpha e^{-\frac{x^2+y^2}{2}} dx dy$

考虑换元: $x = r \sin \theta, y = r \cos \theta, r: 0 \rightarrow +\infty, \theta: 0 \rightarrow 2\pi$ 数学期望还有一大重要的性质: 将事件的示性函数与概率联系在一起. 考虑 $I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A \end{cases}$, 可以发现示性函数实际上是一个取值为 $\{0, 1\}$ 的随机变量. 对它求期望, 有 $\mathbb{E}(I_A) = \mathbb{P}(A)$, 也就是说, 事件 A 的示性函数的期望即事件 A 发生的概率. 由这一条性质, 我们可以得到很多公式和不等式.

定理 4.2 (超几何分布的数学期望)

$X \sim H(n, M, N)$, 其实际模型为 N 个产品, 其中 M 个为次品, 从中任取 n 个. X 为这 n 个中的次品数. 则 $\mathbb{E}X = n \frac{M}{N}$.



证明 定义一系列用以示性的随机变量列. 令 $\xi_i = \begin{cases} 1, i \text{ is bag goods} \\ 0, i \text{ is good goods} \end{cases}, i = 1, 2, \dots, n$. 则 $X = \xi_1 + \xi_2 + \dots + \xi_n$, 由数学期望的线性性质, 得 $\mathbb{E}X = \sum_{i=1}^n \mathbb{E}\xi_i$.

由于不放回的抽取, 有超几何分布的性质可知 $\mathbb{P}(\xi_i = 1) = \frac{M}{N}$, 则 $\mathbb{E}\xi_i = \frac{M}{N}, i = 1, 2, \dots, n$. 于是 $\mathbb{E}X = n \frac{M}{N}$.

命题 4.1 (信封匹配问题)

将 n 封不同信封放入 n 个不同信封, 每封信有且只有一个信封与其匹配, 求正确匹配信的平均数量.



证明 同样地, 采取与上面同样的定义一系列随机变量的方法.

假设 X 为正确匹配的信的数量. 令 $\xi_i = \begin{cases} 1, i \text{ is right} \\ 0, i \text{ is not right} \end{cases}$ 则 $X = \xi_1 + \xi_2 + \dots + \xi_n$, 由数学期望的线性性质: $\mathbb{E}X = \sum_{i=1}^n \mathbb{E}\xi_i$.

而 $\mathbb{P}(\xi_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$. 故而 $\mathbb{E}\xi_i = \frac{1}{n}, i = 1, 2, \dots, n$.

故 $\mathbb{E}X = 1$.

定理 4.3 (Jordan 公式)

设 A_1, \dots, A_n 均为事件, 则

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} \mathbb{P}(A_{j_1} A_{j_2} \dots A_{j_k})$$



证明 设 $I[A_i]$ 表示事件 A_i 的示性函数, 有 $I[\overline{A_i}] = 1 - I[A_i]$ 和 $I[A_i A_j] = I[A_i] I[A_j]$. 于是有

$$I\left[\bigcup_{i=1}^n A_i\right] = 1 - I\left[\bigcap_{i=1}^n \overline{A_i}\right] = 1 - \prod_{i=1}^n I[\overline{A_i}] = 1 - \prod_{i=1}^n (1 - I[A_i]) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} I[A_{j_1} A_{j_2} \dots A_{j_k}]$$

两边同求期望, 由数学期望的线性性质, 得 Jordan 公式.

定理 4.4 (Markov 不等式)

X 为随机变量, 满足 $\mathbb{P}(X \geq 0) = 1$ (即 $X \leq 0$ a.s.), 则 $\forall c > 0, \alpha > 0, \mathbb{P}(X \geq c) \leq \frac{\mathbb{E}X^\alpha}{c^\alpha}$



证明 考虑 $I[X \geq c]$: 当且仅当 $X \geq c$ 时其值为 1.

于是有 $I[X \geq c] \leq 1 \leq \frac{X^\alpha}{c^\alpha}$.

对两边同时求期望, 得 $\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}X^\alpha}{c^\alpha}$.

事实上, 这是一个很松的不等式, 但由于它需要的条件相当少, 它依然成为了概率论中应用最广泛、最重要的不等式之一. 特别地, 令 $\alpha = 2$, 并用 $|X - \mathbb{E}X|$ 替换 X , 得到 Chebyshev 不等式

定理 4.5 (Chebyshev 不等式)

$$\forall c > 0, \mathbb{P}(|X - \mathbb{E}X| \geq c) \leq \frac{\text{Var}X}{c^2}$$



4.3 协方差、相关系数

4.4 条件数学期望

第 5 章 收敛性与中心极限定理

5.1 矩母函数

5.2 特征函数

5.3 多元正态分布

5.4 大数定律

5.5 中心极限定理

第二部分

补充内容

第三部分

习题整理 (含小测题、历年题)

第 6 章 茆书上习题

第 7 章 吴昊讲义上习题

第8章 小测题

8.1 第一次小测

1. 设有 n 次独立重复随机试验, 每次出现 0,1,2 的概率分别是 p_0, p_1, p_2 , 其中 $p_0 + p_1 + p_2 = 1$, 求 n 次试验中 1 和 2 都至少出现一次的概率.
2. 设 B_1, B_2, \dots, B_n 为一系列相互独立的事件 (即它们中任意有限个都是相互独立的). $P(B_n) = p_n, 0 < p_n < 1, \sum_{n=1}^{\infty} p_n = \infty$, 证明:

$$P\left(\bigcup_{n=m}^{\infty} B_n\right) = 1, \forall m$$

3. 考虑进行 n 次独立试验, 第 i 次试验成功的概率为 $\frac{1}{2^{i+1}}$, 令 P_n 表示总的成果次数为奇数的概率.
 - (1). 导出 P_{n-1} 和 P_n 的递推公式.
 - (2). 导出 P_n 的表达式.
4. 假设 E, F 为任意两个事件, $P(F) > 0$, 证明:

$$P(E | E \cup F) \geq P(E | F)$$

5. 证明 $|P(AB) - P(A) - P(B)| \leq \frac{1}{4}$
6. 某赌徒口袋里有一枚均匀硬币和一枚两面都是“正面”的硬币, 他随机地摸出一枚,
 - (1). 假设他抛掷这枚硬币后出现了正面, 求这枚硬币是均匀硬币的概率.
 - (2). 假设他抛掷这枚硬币 k 次出现了 k 次正面, 求这枚硬币是均匀硬币的概率.

8.2 第二次小测

1. 考虑 n 次独立重复抛掷一枚硬币, 每次硬币正面朝上的概率为 p . 证明有偶数个正面朝上的概率为 $\frac{1+(q-p)^n}{2}$. 其中 $q = 1 - p$.

2.(1). 设 X_1, X_2, \dots, X_n 为相互独立的 $N(0, 1)$ 变量, b_1, b_2, \dots, b_n 为实数, 证明: $b_1 X_1 + \dots + b_n X_n \sim N(0, \sigma^2)$, 其中 $\sigma^2 = b_1^2 + \dots + b_n^2$.

(2). 设 $\xi_1, \xi_2, \dots, \xi_6$ 为相互独立的 $N(0, 1)$ 变量. 记

$$\eta = \frac{\xi_1 \xi_4 + \xi_2 \xi_5 + \xi_3 \xi_6}{\sqrt{\xi_4^2 + \xi_5^2 + \xi_6^2}}$$

求 η 的分布.

3. 设 ξ 和 η 独立, $\eta \sim \gamma(\lambda, a + b)$, $\xi \sim \beta(a, b)$, 求 $\xi\eta$ 的分布.

4. 设 X_1, \dots, X_n 为相互独立同分布的随机变量, 均服从参数为 $\lambda > 0$ 的指数分布, N 服从参数为 β 的 Poisson 分布, 且与 X_1, X_2, \dots 独立. 记 $X_0 = 0$, 定义 T 如下: 对每个 ω , $T(\omega) = \max_{0 \leq k \leq N(\omega)} X_k(\omega)$.

(1). 证明 T 是随机变量.

(2). 求 T 的分布函数与密度函数.

5. 设 X_1, \dots, X_n 为相互独立的标准正态随机变量, 令 $S_j = X_1 + \dots + X_j$, $1 \leq j \leq n$.

(1). 求在给定 $S_j = x$ 的条件下, S_n 的条件分布 ($1 \leq j < n$)

(2). 求在给定 $S_n = y$ 的条件下, S_j 的条件分布 ($1 \leq j < n$).

8.3 第三次小测

1. 设 X 服从参数为 M, N 和 n 的超几何分布, 即 $\mathbb{P} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, k = 0, 1, 2, \dots$, 其中 $n \leq M \leq N$. 求 EX 和 $VarX$
2. 设 X 和 Y 为相互独立的标准正态随机变量, 求 XY 的特征函数.
3. 设 X_1, \dots, X_n, \dots 是一列相互独立同分布的标准正态变量, N_n 服从二项分布 $B(n, p)$ ($0 < p < 1$) 并与 X_1, \dots, X_n, \dots 独立. 证明:

$$\xi_n = \frac{\sum_{k=1}^{N_n} X_k}{\sqrt{n}} \xrightarrow{d} N(0, p), n \rightarrow \infty$$

4. $g(t)$ 为定义在 $[0, +\infty)$ 上的单调递减非负连续函数, 记 $\int_0^{+\infty} g(t)dt = a, \int_0^{+\infty} tg(t)dt = b$, 设 $0 < a, b < +\infty$. 设随机变量 X, Y 的联合密度函数为 $p(x, y) = g(x^2 + y^2), -\infty < x, y < +\infty$.

- (1) 求 a 的值.
- (2) 求随机变量 X, Y 的期望、方差.
- (3) 求 X, Y 的相关系数.
- (4) 设 X, Y 相互独立, 证明 X, Y 均服从正态分布.

第9章 历年卷

9.1 2023-2024 概率论 3.5 学分 zlx 班

1. 证明:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \cdots + (-1)^n P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

2. 两批零件, 第一种 n_1 个, 寿命 $X_1, \cdots, X_{n_1} \sim E(\lambda_1)$, 第二种 n_2 个, 寿命 $Y_1, \cdots, Y_{n_2} \sim E(\lambda_2)$, 有一个零件失效则失效. 记 T 为失效时间.

(1). 证明: $T \sim E(\lambda_1 n_1 + \lambda_2 n_2)$;

(2). 求第一种零件失效导致失效的概率.

3. X_1, \cdots, X_n 独立同分布 $\sim P(\lambda)$. 且有

$$S = X_1 + \cdots + X_n, \bar{X} = \frac{X_1 + \cdots + X_n}{n}, T = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

(1). 计算 ET ;

(2). $S = s (s = 0, 1, \cdots)$ 时, 证明 $X_i \sim B(s, \frac{1}{n})$;

(3). 计算 $E(T | S)$.

4. $p(x, y) = C(x-y)^2 e^{-\frac{1}{2}(x^2+y^2)}$.

(1). 求 C ;

(2). 求 $P_X(x), P_Y(y)$;

(3). 计算 r_{XY} ;

(4). 证明 $X+Y, X-Y$ 独立.

5. X_i 独立同分布, $E|X_1| < \infty, \mu = EX_1$, 证明:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

6. N_p 服从几何分布, 参数为 p, X_i 独立同分布 $\sim N(\mu, \sigma^2), Y_p = \sum_{k=1}^{N_p} X_k$.

(1). 证明 Y_p 是随机变量;

(2). 求 EY_p ;

(3). 求 $VarY_p$;

(4). 证明 Y_p 不是正态随机变量.

7. (附加题)(1). 证明 Y_p 是连续型随机变量;

(2). $p \rightarrow 0$, 证明 pY_p 收敛到某个分布函数.

9.2 2022-2023 概率论 3.5 学分 zlx 班

1. 设事件序列 $\{A_n\}$ 相互独立, $P(A_n) = \frac{1}{n^2}$, 求 $P(\bigcup_{n=m}^{\infty} A_n)$.
2. $\{\xi_k\}$ 为独立同分布的随机变量, $\xi_k \sim E(\lambda)$, $S_k = \sum_{i=1}^k \xi_i$.
 - (1). 求证 S_k 服从参数为 k 和 λ 的 gamma 分布.
 - (2). 令 $N = \max_k \{S_k \leq x\}$. 求证: N 服从参数为 λ_x 的 Poisson 分布. (提示: $P(N = k) = p(S_k \leq x, S_{k+1} > x)$)
3. X_1, X_2 为标准正态随机变量, $Y = e^{X_1 - X_2}$.
 - (1). 求 $EY, VarY$.
 - (2). 求 Y 与 X_1 的协方差 $Cov(Y, X_1)$ 以及相关系数 β .
 - (3). 设随机变量 $Z = Y - \alpha X_1 - \beta X_2$. 求 α, β 使得 Z 与 X_1 和 X_2 均不相关.
4. X, Y 为随机变量, 密度函数

$$p(x, y) = c(x + y) \exp -x^2 - y^2$$

- (1). 求证: $X + Y$ 与 $X - Y$ 相互独立.
- (2). 求 c 的值.
- (3). 求 $X + Y$ 与 $X - Y$ 的密度函数.
5. $\{\xi_n\}$ 为一列独立同分布随机变量, $E\xi_1 = \mu, Var\xi_1 = \sigma^2, T_n = \sum_{1 \leq i < j \leq n} \xi_i \xi_j$.
 - (1). 求 T_n 的期望、方差;
 - (2). 求证存在常数 c 使得 $\frac{T_n}{n^2} \xrightarrow{P} c$ 并求 c 的值.
6. $\{\xi_i\}$ 为一列独立同分布随机变量, $E\xi_i = \mu, N_p$ 为服从几何分布的随机变量 ($P(N_p = k) = p(1-p)^k$). $Y_p = \sum_{i=1}^{N_p} \xi_i$.
 - (1). 求证: Y_p 为随机变量.
 - (2). 求证: $p \rightarrow 0$ 时, pY_p 依分布收敛, 并求其收敛分布.

9.3 2021-2022 概率论 3.5 学分 zlx 班

- (1). $\{A_n\}$ 为独立事件列, $\sum_{n=1}^{\infty} P(A_n) = \infty$, 求证: $P(\bigcup_{n=m}^{\infty} A_n) = 1, \forall m \in \mathbb{N}_+$.
- (2). X, Y 为独立且服从几何分布的随机变量, 参数为 p , 求证: $P(X = i | X + Y = n) = \frac{1}{n-1}, i = 1, 2, \dots$
2. $S_n = \sum_{k=1}^n \sin^2(kU), U \sim U(0, 2\pi)$.
 - (1). 求 ES_n .
 - (2). 证明 $\frac{S_n}{n} \xrightarrow{P} c$ 并求出 c .
3. $X, Y \sim N(0, 0, 1, 1, r)$.
 - (1). 求出 a 使得 Y 与 $X - aY$ 不相关, 并求出 $X - aY$ 的分布函数.
 - (2). 求 X^2 与 Y^2 的相关系数.
4. $p(x, y) = C(x - y)^2 \exp\{-\frac{1}{2}(x^2 + y^2)\}$ 为 X, Y 联合密度.
 - (1). 求 C .
 - (2). 求 $EX, EY, VarX, VarY$.
 - (3). 证明: $X + Y$ 与 $X - Y$ 独立.
 - (4). 求 $X^2 + Y^2$ 的密度.
5. 针对 X, Y 为离散形式的情况, 利用期望的定义证明:

$$EXY = EX \cdot EY$$

6. 已知 $\xi_i \xrightarrow{P} \varepsilon(\lambda), E(\xi_1) = \frac{1}{\lambda}, N_n \sim \mathcal{P}(n)$ (参数为 n 的 Poisson 分布). 记 $\eta_n = \frac{\sum_{k=1}^{N_n} \xi_k}{n}$.
 - (1). 证明 η_n 是随机变量, 求出 η_n 的特征函数.
 - (2). $\sqrt{n}(\eta_n - a) \xrightarrow{d} N(0, b)$ 并求出对应的 a 和 b .