

# 《随机过程引论》

♡ 2023 -2024 学年 第一学期 ♡



## 课程信息及要求

---

- ◇ 教师信息: 何萍, pinghe@mail.shufe.edu.cn  
办公室: 红瓦楼 815 室, 电话: 021-65908857
- ◇ 学习要求: 不迟到不早退不旷课, 按时交作业 (每周二课前)
- ◇ 考核方式: 平时成绩 (30%) + 期末成绩 (70%)



## 课程内容

- ◇ 随机过程, 是与概率论一样的基础数学, 都起源于现实生活中的问题. 比如, 赌博

其结构  $\left\{ \begin{array}{l} \text{在古代是简单, 有限的(如, 机会性游戏),} \\ \text{在现代是综合, 无限的(如, 全球股市).} \end{array} \right.$

- ◇ 测度论不是我们的先修课程.  
随机过程的概念很广泛, 研究内容几乎包括概率论的全部.  
本课程介绍一些概率论的基础知识之后,

基本架构按随机过程的模型来分章节(见书中目录).



## 参考书

---

- ① Feller, W., Probability Theory and its Application,  
Vol. I (1959: Third edition),  
Vol. II (1970), Wiley & Son
- ② S.M.劳斯(何声武等译), 随机过程, 中国统计出版社, 1997
- ③ Ross, S.M., Introduction to Probability Models (12th edition),  
Academic Press, San Diego, 2019



# 第一章 概率论述要



## 1 随机变量

- 概率空间
- 随机变量及其分布
- 期望与方差

## 2 随机向量

- 联合分布
- 协方差与协方差矩阵
- 例子: Gauss 分布
- 随机向量函数: 独立和, 顺序统计量

## 3 极限定理

- (弱)大数定律
- Borel-Cantelli 引理
- Borel 强大数定律及其应用: 投资陷阱
- 极限与期望交换

## 4 矩母函数与母函数

- 矩母函数
- 母函数



## 继续使用概率论的基本概念与符号.

- $(\Omega, \mathcal{F})$  为可测空间, 其中
  - 样本空间  $\Omega$ : 一个非空集合,
  - $\Omega$  上的事件域  $\mathcal{F}$ : 事件的全体, 满足
    - (1)  $\emptyset, \Omega \in \mathcal{F}$ ;
    - (2)  $A \in \mathcal{F}$  蕴含  $A^c \in \mathcal{F}$ ;
    - (3)  $A_n \in \mathcal{F}, n \geq 1$  蕴含  $\bigcup_{n \geq 1} A_n \in \mathcal{F}$ .



## 公理化概率

- 可测空间  $(\Omega, \mathcal{F})$  上的概率测度  $\mathbb{P}(\cdot): \mathcal{F} \rightarrow [0, 1]$ , 满足
  - (1) 非负性:  $\mathbb{P}(A) \geq 0, A \in \mathcal{F}$ ;
  - (2) 规范性/正则性:  $\mathbb{P}(\Omega) = 1$ ;
  - (3) 可列可加性: 对  $\mathcal{F}$  中互斥的可列个事件  $\{A_n: n \geq 1\}$ , 有

$$\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mathbb{P}(A_n).$$

三元组  $(\Omega, \mathcal{F}, \mathbb{P})$  称为概率空间.



## 一些重要的性质

- $\mathbb{P}$  是  $(\Omega, \mathcal{F})$  上满足  $\mathbb{P}(\Omega) = 1$  的非负集合函数.

(李贤平, 1996: 定理1.5.1)

$\mathbb{P}$  具有可列可加性  $\Leftrightarrow \mathbb{P}$  具有  $\begin{cases} \text{有限可加性,} \\ \text{下连续性.} \end{cases}$

(下连续性:) 若  $A_n \in \mathcal{F}, n \geq 1$  且递增, 则

$$\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

- 对于一般的测度,

$\mathbb{P}$  具有可列可加性  $\Leftrightarrow \mathbb{P}$  具有  $\begin{cases} \text{有限可加性,} \\ \text{次可列可加性.} \end{cases}$



## 概率空间的例

例 1. 设  $\Omega$  是非空集合,  $A \subset \Omega$ : 非空子集. 定义

$$\mathcal{F} = \{\emptyset, \Omega, A, \bar{A}\},$$

$$\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1, \mathbb{P}(A) = p, \mathbb{P}(\bar{A}) = 1 - p,$$

$$0 < p < 1.$$

那么,  $(\Omega, \mathcal{F}, \mathbb{P})$  是最简单的非平凡的概率空间. #



## 例 2. (古典概率模型)

设  $\Omega$  是一个有限非空集合,  $\mathcal{F}$  是  $\Omega$  的子集全体,  
对  $A \subset \Omega$ , 定义

$$\mathbb{P}(A) := \frac{\#(A)}{\#(\Omega)},$$

其中,  $\#$  表示  $\Omega$  上的计数测度.

$(\Omega, \mathcal{F}, \mathbb{P})$  是一个古典概率空间.

等价的, 每个基本结果等可能地发生.

#



用 Lebesgue 测度给出的等可能性.

例 3. (几何概率空间)

设  $\Omega = (a, b)$ ,  $\mathcal{F}$  是由  $\Omega$  中所有开集生成的  $\sigma$ -代数.  
令  $m$  是  $(\Omega, \mathcal{F})$  的 Lebesgue 测度, 比如,

$$\text{对 } (c, d) \subseteq (a, b), \quad m(c, d) = d - c.$$

对任意  $A \in \mathcal{F}$ ,

$$\mathbb{P}(A) = \frac{m(A)}{b - a}.$$

那么,  $(\Omega, \mathcal{F}, \mathbb{P})$  是一个几何概率空间.

#



## 随机变量=实值可测函数

注. 设  $(\Omega, \mathcal{F})$  和  $(\Omega', \mathcal{F}')$  是两个给定可测空间,  
 $\xi: \Omega \mapsto \Omega'$  上映射. 若

$$\xi^{-1}(A) \in \mathcal{F}, \forall A \in \mathcal{F}',$$

则称  $\xi$  是可测映射. 若  $(\Omega', \mathcal{F}') = (\mathbf{R}, \mathcal{B})$  则为随机变量:

给定概率空间:  $(\Omega, \mathcal{F}, \mathbb{P})$ .

### 定义 1.1.3:

设  $\xi: \Omega \longrightarrow \mathbf{R}$ , 如果对任何  $x \in \mathbf{R}$ ,

有  $\{\xi \leq x\} \in \mathcal{F}$ , 那么  $\xi$  称为随机变量.



## 常用最小 $\sigma$ -代数

- 设  $\xi: \Omega \rightarrow \mathbf{R}$  上的映射,

称  $\sigma(\xi) := \xi^{-1}(\mathcal{B})$  为由  $\xi$  生成的  $\sigma$ -代数.

- 更一般的, 设  $T$  是一个指标集, 对  $i \in T$ , 有  $\xi_i: \Omega \rightarrow \mathbf{R}$ .  
记

$$\sigma(\xi_i, i \in T) := \sigma\{\xi_i^{-1}(\mathcal{B}), i \in T\}$$

(: 由  $\{\xi_i^{-1}(\mathcal{B}), i \in T\}$  生成的  $\sigma$ -代数)

称为由  $\{\xi_i, i \in T\}$  生成的  $\sigma$ -代数.



例 4. 班上有 6 个学生, 编号分别是 1 ~ 6 号, 其中

1 ~ 4 号是男生, 5 ~ 6 号是女生,  
1, 2, 6 号戴眼镜, 3, 4, 5 号不戴眼镜.

任取一名学生,

$$X = \begin{cases} 0, & \text{女生,} \\ 1, & \text{男生,} \end{cases} \quad Y = \begin{cases} 0, & \text{不戴眼镜,} \\ 1, & \text{戴眼镜,} \end{cases}$$

- (1) 写出概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$ ;
- (2) 写出  $\sigma(X), \sigma(Y), \sigma(X, Y)$ ;
- (3) 计算  $\mathbb{P}(X = 0, Y = 1)$ .



### 定义 1.1.3:

相应的, 称函数  $F(x) := \mathbb{P}(\xi \leq x), x \in \mathbf{R}$  为  $\xi$  的分布函数,

记为  $\xi \sim F(x)$ .

分布函数的性质有

- (1)  $F$  是递增的;
- (2)  $F$  是右连续的;
- (3) (规范性)  $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$ .

一般地,  $\mathbf{R}$  上满足上面性质的函数称为是分布函数, 简称分布.

(概率论中, 概率测度通常也称为分布, 不刻意区分这两个名词.)



## 常见分布

- 参数为  $\lambda > 0$  的 Poisson 分布  $P(\lambda)$ :

$$\mathbb{P}(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

常用来描述某段时间内某随机发生的事件发生的次数.

注. Poisson 分布是许多分布的近似.



例 1.1.2 : (配对问题) 将  $n$  个标号的球放入  $n$  个标号的盒子里, 假设每个盒子放一个球. 用  $\zeta$  表示  $n$  组中形成的配对数, 那么  $\zeta$  的分布

$$p_k(n) := \mathbb{P}(\zeta = k) = \frac{1}{k!} \sum_{j=0}^{n-k} (-1)^j \frac{1}{j!}, \quad 0 \leq k \leq n. \quad (1)$$

当  $n$  很大时, 有

$$\mathbb{P}(\zeta = k) \simeq \frac{e^{-1}}{k!}, \quad (2)$$

也就是说  $\zeta$  差不多服从参数为 1 的 Poisson 分布.



证. (1) 用  $A_i$  表示第  $i$  组配成对, 有

$$\mathbb{P}(A_i) = \frac{(n-1)!}{n!}, \quad \mathbb{P}(A_i A_j) = \frac{(n-2)!}{n!}, \quad i \neq j, \dots,$$

$$\mathbb{P}(A_1 A_2 \cdots A_n) = \frac{1}{n!}.$$

由加法公式, 对任意  $n \geq 1$ ,

$$\begin{aligned} p_0(n) &= 1 - \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \\ &= 1 - [n\mathbb{P}(A_1) - C_n^2\mathbb{P}(A_1 A_2) + \cdots + (-1)^{n-1}\mathbb{P}(A_1 \cap \cdots A_n)] \\ &= 1 - 1 + \frac{1}{2!} + \cdots + (-1)^n \frac{1}{n!} = \sum_{j=0}^n (-1)^j \frac{1}{j!}. \end{aligned}$$



而根据乘法原理, 对任意  $k \leq n$ ,

$$p_k(n) = \frac{C_n^k p_0(n-k) \cdot (n-k)!}{n!} = \frac{p_0(n-k)}{k!}.$$

所以

$$\mathbb{P}(\xi = k) = \frac{1}{k!} \sum_{j=0}^{n-k} (-1)^j \frac{1}{j!}, \quad 0 \leq k \leq n.$$

(2) 取  $n$  充分大, 有

$$\mathbb{P}(\xi = k) \approx \frac{e^{-1}}{k!},$$

也就是说,  $\xi$  差不多是参数是 1 的 Poisson 分布.

#



- 参数为  $\alpha(>0)$  的指数分布  $\text{Exp}(\alpha)$ : 密度函数

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

对应的分布函数

$$F(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

(遗忘性:) 设  $\zeta \sim \text{Exp}(\alpha)$ , 对任何  $x, y > 0$ , 有

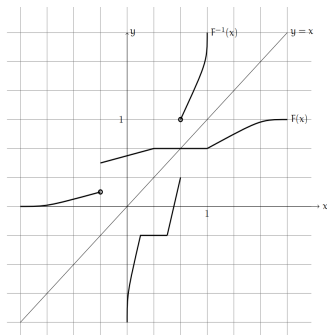
$$\mathbb{P}(\zeta > x + y | \zeta > x) = \mathbb{P}(\zeta > y).$$



## 随机变量函数

习题: 3. (如何用伪随机数生成随机数.)

如果随机变量  $U \sim U(0, 1)$ ,  $F(x)$  是一个给定的分布函数, 则  $F^{-1}(U)$  的分布函数是  $F$ . 其中  $F^{-1}$  是  $F$  的(广义)反函数.



## 随机变量的数学期望

- 当  $\zeta$  是离散型时, 若下面级数存在,

$$\mathbb{E}[\zeta] = \sum_{n=1}^{\infty} x_n \mathbb{P}(\zeta = x_n);$$

- 当  $\zeta$  是连续型时, 若下面积分存在,

$$\mathbb{E}[\zeta] = \int_{-\infty}^{\infty} xp(x)dx.$$



## 一般随机变量的数学期望

设  $\xi \sim F(x)$ , 如果

$$\int_{-\infty}^{\infty} |x| dF(x) < \infty,$$

其中积分是 Riemann-Stieltjes 积分, 那么  $\xi$  的数学期望是

$$\mathbb{E}[\xi] = \int_{-\infty}^{\infty} x dF(x).$$

期望是随机变量的中心, 或者平均.



## 期望的性质:

- ①  $\mathbb{E}[1] = 1;$
- ②  $\mathbb{E}[a\zeta + b\eta] = a\mathbb{E}[\zeta] + b\mathbb{E}[\eta];$
- ③ 若  $\zeta \sim F(x)$ ,  $f: \mathbf{R} \mapsto \mathbf{R}$  是 Borel 可测函数, 则

$$\mathbb{E}[f(\zeta)] = \int_{-\infty}^{\infty} f(x) dF(x),$$

若上式两边存在.

- ④ 如果  $\zeta, \eta$  独立, 那么  $\mathbb{E}[\zeta\eta] = \mathbb{E}[\zeta] \cdot \mathbb{E}[\eta].$



## 随机变量的方差

对于平方可积的随机变量, 可定义方差

$$D[\xi] := \mathbb{E}[(\xi - \mathbb{E}[\xi])^2] = \mathbb{E}[\xi^2] - (\mathbb{E}[\xi])^2.$$

方差描述随机变量的随机性大小.

● 性质:

- ①  $D[1] = 0$ ;
- ②  $D[c\xi] = c^2 D[\xi]$ ;
- ③ 如果  $\xi, \eta$  独立, 那么  $D[\xi + \eta] = D[\xi] + D[\eta]$ .



## 方差的性质(续)

对于随机变量  $\xi$ , 假设  $\mathbb{E}\xi^2 < \infty$ .

- 若用一个常数  $c$  估计  $\xi$ , 试问  $c$  的值为何值时,

均方误差  $\mathbb{E}[(\xi - c)^2]$  最小?

事实上, 由

$$\mathbb{E}[(\xi - c)^2] = D(\xi) + (\mathbb{E}[\xi] - c)^2,$$

可知

当且仅当  $c = \mathbb{E}\xi$  时,  $\mathbb{E}[(\xi - c)^2]$  最小.



## 随机向量的联合分布

给定概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$ . 设  $X_1, \dots, X_n$  为其上  $n$  个随机变量, 则

$(X_1, \dots, X_n)$  为随机向量.

相应的, 对任意实数  $x_1, x_2, \dots, x_n$ , 定义联合分布函数

$$F(x_1, x_2, \dots, x_n) := \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n).$$

对应的, 每一个分量  $X_i$  的分布函数  $F_i(x_i)$  称为边缘分布函数.



## 随机向量的联合分布

- 如果  $(X_1, \dots, X_n)$  是离散型, 其联合分布通常用分布律刻画:

$$p(x_1, \dots, x_n) := \mathbb{P}(X_1 = x_1, \dots, X_n = x_n),$$

$$\forall x_1, \dots, x_n \in \mathbf{R};$$

- 如果联合分布函数有密度:

$$dF(x_1, \dots, x_n) = p(x_1, \dots, x_n) dx_1 \cdots dx_n,$$

就说随机向量是连续型的, 相应密度称为联合密度函数.



## 独立性

- 若对任意实数  $x_1, x_2, \dots, x_n$ ,

$$F(x_1, x_2, \dots, x_n) = F_1(x_1)F_2(x_2) \cdots F_n(x_n),$$

则称随机变量  $X_1, \dots, X_n$  相互独立. 另有如下充要条件:

- (连续型) 对任意实数  $x_1, x_2, \dots, x_n$ ,

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

其中  $p_i$  是  $X_i$  的密度函数.

- (离散型) 对任意实数  $x_1, x_2, \dots, x_n$ ,

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_n = x_n).$$



## 协方差与相关系数

针对平方可积的二维随机向量  $(X, Y)$ ,

- $X, Y$  的协方差

$$\text{cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])];$$

- $X, Y$  的相关系数

$$\rho(X, Y) := \frac{\text{cov}(X, Y)}{\sqrt{D(X)D(Y)}}.$$

(二者一定程度上描述  $X, Y$  之间的线性关系.)



注. 1. Cauchy-Schwarz 不等式:

$$|\text{cov}(X, Y)| \leq \sqrt{D(X)D(Y)};$$

2. 相关系数的绝对值不超过 1.



## 协方差矩阵

针对  $n$  维随机向量  $(X_1, \dots, X_n)$ , 由协方差组成的矩阵

$$[\text{cov}(X_i, X_j)]_{1 \leq i, j \leq n}$$

称为是随机向量的协方差矩阵.

**注.** 协方差矩阵是一个对称非负定矩阵.

事实上, 对任意  $(x_1, \dots, x_n) \in \mathbf{R}^n$ ,

$$\sum_{i,j} x_i \text{cov}(X_i, X_j) x_j = D \left( \sum_{i=1}^n x_i X_i \right) \geq 0.$$



## 注. 协方差矩阵

- 当且仅当满秩时, 它是正定的.
- 当它不是满秩时, 就说它是退化的. 比如, 当  $n = 2$  时,

协方差矩阵退化等价于相关系数  $|\rho| = 1$ ,

即随机变量线性相关; 一般场合类似.



例 1.2.1 设  $A = (a_{i,j})$  是对称正定矩阵, 密度函数

$$p(x) = \frac{1}{\sqrt{|A|}(\sqrt{2\pi})^n} \exp \left\{ -\frac{1}{2} x A^{-1} x^T \right\}, \quad x \in \mathbf{R}^n.$$

(后面称为中心化正态分布)

注.  $A$  对称正定当且仅当  $A^{-1}$  对称正定.

- 在上式使用  $A^{-1}$  的目的只是为了与  $n=1$  时的表示和谐.
- 存在对称正定矩阵  $B = (b_{i,j}) : B^2 = A$ , 从而

$$A^{-1} = (B^{-1})^2.$$



## $n$ 维正态分布, 亦称 Gauss 分布

### 定义:

如果  $a = (a_1, a_2, \dots, a_n) \in \mathbf{R}^n$ ,  $A$  对称正定, 那么密度函数

$$p(x) = \frac{1}{\sqrt{|A|}(\sqrt{2\pi})^n} \exp\left\{-\frac{1}{2}(x-a)A^{-1}(x^T - a^T)\right\}, \quad x \in \mathbf{R}^n,$$

对应的分布称为是 (参数为  $a, A$  的)  $n$  维正态分布. 记为

$$X = (X_1, X_2, \dots, X_n) \sim N(a, A).$$

- 当  $a = 0$  时称为中心化正态分布; 进一步地,  $A$  是单位矩阵  $I$ , 称为是标准正态分布, 记为  $N(0, I)$ .
- 若  $X \sim N(a, A)$ , 则

$$(X - a)\sqrt{A}^{-1} = (X - a)B^{-1} \sim N(0, I).$$



注. 一个随机向量是正态分布的当且仅当

其密度函数为  $x \mapsto C \cdot \exp(-\phi(x))$ ,

其中  $C$ : 常数,  $\phi$ : 有最小值的二次型.



## Gauss 分布的性质

命题:

设  $X = (X_1, \dots, X_n) \sim N(a, A)$ , 则

- 其每个分量都是正态分布的. 更一般地,
- $\zeta$  的任何多个分量组成的随机向量仍是正态分布的.

证. 只证  $(X_2, \dots, X_n)$  是正态分布.



证. 只证  $(X_2, \dots, X_n)$  是正态分布. 首先

$$(x_2, \dots, x_n) \mapsto \int_{-\infty}^{+\infty} p(x_1, x_2, \dots, x_n) dx_1,$$

不妨设  $a = 0$ , 则二次型可写为

$$xA^{-1}x^T = a_{11}(x_1 + \sum_{i=2}^n b_i x_i)^2 + x_0 A_0 x_0^T,$$

其中  $x_0 = (x_2, \dots, x_n)$ ,  $A_0$  是  $n-1$  阶对称正定方阵. 所以

$$\int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{2} x A^{-1} x^T \right\} dx_1 = C \cdot \exp \left\{ -\frac{1}{2} x_0 A_0 x_0^T \right\}, \quad C: \text{常数},$$

故  $(X_2, \dots, X_n)$  也是正态分布的.



## Gauss 分布的期望

- 如果  $X$  是中心化的, 那么

$$\mathbb{E}X_i = \int_{\mathbf{R}^n} x_i p(x) dx = 0, \quad x = (x_1, x_2, \dots, x_n),$$

因为  $x_i f_X(x)$  是中心对称的.

- 如果  $X \sim N(a, A)$ , 那么  $X - a$  是中心化的, 所以

$$\mathbb{E}X_i = a_i, \quad i = 1, 2, \dots, n.$$



## Gauss 分布的协方差阵

- 不妨设  $X$  是中心化的, 因为  $X - a$  和  $X$  有相同协方差.

- 当  $X$  是标准正态时, 其协方差矩阵就是单位矩阵; 此时,

$X_1, \dots, X_n$  相互独立且标准正态分布的.

- (用矩阵的语言)

如果  $Y = XB^{-1}$  是标准正态分布, 则其协方差矩阵是单位矩阵, 从而

$$\mathbb{E}X^T X = B \cdot \mathbb{E}Y^T Y \cdot B = B^2 = A.$$



## 随机变量的独立和

### 卷积公式

如果  $X, Y$  独立且密度函数分别为  $f$  和  $g$ , 则它们的和  $X + Y$  的密度函数  $h$  是它们的卷积:

$$h = f * g,$$

i.e., 对于  $x \in \mathbf{R}$ ,

$$h(x) = \int_{-\infty}^{+\infty} f(x-y)g(y)dy.$$

▷ 有再生性的常见分布:

二项分布, Poisson 分布, 正态分布, Gamma 分布, 等等.



例 1.2.5 设  $\{X_n\}$  是独立同分布随机序列, 服从参数为  $\alpha$  的指数分布. 令  $S_0 = 0$ ,

$$S_n = \sum_{k=1}^n X_k.$$

若视  $X_n$  为等待时间,  $S_n$  可视为第  $n$  个随机信号到达的时间, 自然  $S_n$  服从  $\Gamma$  分布  $\Gamma(n, \alpha)$ .

对任何  $t \geq 0$ , 令

$$N(t) := \sup\{n : S_n \leq t\},$$

计算  $N(t)$  的分布.

注.

$$S_n \leq t < S_{n+1} \text{ 当且仅当 } N(t) = n.$$

也就是说,  $N(t)$  是  $t$  时刻前已到达的信号数量.



解.

$$\begin{aligned}
 \mathbb{P}(N(t) = n) &= \mathbb{P}(S_n \leq t < S_{n+1}) \\
 &= \mathbb{P}(S_n \leq t < S_n + X_{n+1}) \\
 &= \int_0^t \frac{\alpha^n}{(n-1)!} x^{n-1} e^{-\alpha x} dx \int_{t-x}^{\infty} \alpha e^{-\alpha y} dy \\
 &= e^{-\alpha t} \frac{\alpha^n}{(n-1)!} \int_0^t x^{n-1} dx = e^{-\alpha t} \frac{(\alpha t)^n}{n!},
 \end{aligned}$$

也就是说,  $N(t)$  服从参数为  $\alpha t$  的 Poisson 分布.

#



## 顺序统计量

注. 考虑独立随机序列  $\{\xi_i\}_{i=1}^n$  都是连续型分布的, 可认为它们互不相同, 重新按从小到大的顺序排列为

$$\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(n)}.$$

定义:

对于任意的  $\omega \in \Omega$ ,  $\xi_{(i)}(\omega)$  表示  $n$  个实数  $\xi_1(\omega), \dots, \xi_n(\omega)$  从小到大排列时的第  $i$  个数, 即符号  $(i)$  代表随机序号:  
对任意  $\omega$ ,  $(i)(\omega) = j$  当且仅当

数  $\xi_j(\omega)$  在以上排列中排在第  $i$  个位置.

随机变量  $\xi_{(1)}, \dots, \xi_{(n)}$  称为  $\{\xi_i\}$  的顺序统计量.



注. (对称性) 说随机向量  $(\xi_1, \dots, \xi_n)$  是**对称的**<sup>+</sup>,  
如果对于  $(1, \dots, n)$  的任意一个置换  $(\sigma_1, \dots, \sigma_n)$ ,

$$(\xi_{\sigma_1}, \dots, \xi_{\sigma_n}) \text{ 与 } (\xi_1, \dots, \xi_n) \text{ 同分布,}$$

也就是说分布与顺序无关.

- ⇒
- ① 由对称性,  $\{\xi_i\}$  的各种不同顺序排列应是等可能发生的.
  - ②  $((1), (2), \dots, (n))$  也是随机向量.  $((1), (2), \dots, (n))$

在  $1, 2, \dots, n$  的所有顺序的集合上是均匀分布的.



(亦可用函数来表达这种分布的对称性: )

说  $(\xi_1, \dots, \xi_n)$  是对称的是指, 对任意非负或有界 Borel 可测函数  $f: \mathbf{R}^n \mapsto \mathbf{R}$ , 有

$$\mathbb{E}f(\xi_1, \dots, \xi_n) = \mathbb{E}f(\xi_{\sigma_1}, \dots, \xi_{\sigma_n}).$$

命题:

若  $\xi_1, \dots, \xi_n$  是独立同分布的, 则它作为随机向量是对称的.



例. 求  $\xi_{(i)}$  的分布,  $i = 2, \dots, n-1$ .

解. 对任何  $x \in \mathbf{R}$ ,  $\xi_{(i)} \leq x$  当且仅当

$\{\xi_i\}_{i=1}^n$  中至少有  $i$  个随机变量不超过  $x$ .

记事件  $A_k :=$  “随机序列中恰有  $k$  个不超过  $x$ ”, 则

$$\begin{aligned}\mathbb{P}(A_k) &= C_n^k \mathbb{P}(\xi_1 \leq x, \dots, \xi_k \leq x, \xi_{k+1} > x, \dots, \xi_n > x) \\ &\quad (\text{利用对称性}) \\ &= C_n^k F(x)^k (1 - F(x))^{n-k} \quad (: \text{二项分布 } B(n, F(x))).\end{aligned}$$

因此所求分布函数是

$$\begin{aligned}\mathbb{P}(\xi_{(i)} \leq x) &= \sum_{k=i}^n C_n^k F(x)^k (1 - F(x))^{n-k} \\ &= \frac{n!}{(i-1)!(n-i)!} \int_0^{F(x)} t^{i-1} (1-t)^{n-i} dt \quad (\text{见习题 10}).\end{aligned}$$



## 一些不等式

### Chebyshev 不等式

对于平方可积的  $X$ ,

$$\mathbb{P}(|X| > \epsilon) \leq \frac{\mathbb{E}[|X|^2]}{\epsilon^2}.$$

仅知道概率分布的二阶矩, 即可由此不等式得到概率的界.

**注.** 常作为理论证明的工具. 比如,

- ① 证明方差为 0 的随机变量就是退化分布的;
- ② 证明 Bernoulli 大数定律.



## Bernoulli 大数定律

对任何  $\varepsilon > 0$ , 有

$$\lim_n \mathbb{P} \left( \left| \frac{\zeta_n}{n} - p \right| > \varepsilon \right) = 0.$$

### 依概率收敛性

设  $\{\zeta_n\}$  是一个随机变量序列,  $\zeta$  是一个随机变量.  
若对任何  $\epsilon > 0$ ,

$$\lim_n \mathbb{P}(\{|\zeta_n - \zeta| \geq \epsilon\}) = 0,$$

则称  $\{\zeta_n\}$  依概率收敛于  $\zeta$ , 记为  $\zeta_n \xrightarrow{P} \zeta$ .



## 几乎处处收敛

设  $\{\xi_n\}$  是一个随机变量序列,  $\xi$  是一个随机变量. 若

$$\mathbb{P}(\lim_n \xi_n = \xi) = 1,$$

则称  $\xi_n$  几乎处处或者以概率 1 收敛于  $\xi$ . 记为  $\xi_n \xrightarrow{\text{a.s.}} \xi$ .

(下面关注两种收敛性的关系.)



注. 对于事件列  $A_n \in \mathcal{F}, n \geq 1$ , 定义上、下极限事件:

- $$\limsup_{n \rightarrow \infty} A_n := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$$

$$= \{\omega \in \Omega : \exists \{k_n\} \subset \mathbb{N} \text{ s.t. } \omega \in A_{k_n}, n \geq 1\};$$

(: 无穷多个  $A_n$  发生.)
- $$\liminf_{n \rightarrow \infty} A_n := \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$$

$$= \{\omega \in \Omega : \exists N \in \mathbb{N} \text{ s.t. } \omega \in A_n, n \geq N\}.$$

(: 至多有限个  $A_n$  不发生.)



## Borel-Cantelli 引理

### 定理 1.3.2:

设  $\{A_n, n \geq 1\}$  是事件列.

(1) 若  $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$ , 则

$$\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0;$$

(2) 若  $\{A_n\}$  是独立事件列且  $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ , 则

$$\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 1.$$



证. (1) 首先

$$\mathbb{P}(\limsup_n A_n) = \lim_n \mathbb{P}(\bigcup_{k \geq n} A_k),$$

因为级数  $\sum_{n=1}^{\infty} \mathbb{P}(A_n)$  收敛, 所以

$$\mathbb{P}(\bigcup_{k \geq n} A_k) \leq \sum_{k \geq n} \mathbb{P}(A_k) \longrightarrow 0.$$



证. (2) 对  $n < N$ , 由于  $\{A_n\}$  独立,

$$\begin{aligned}\mathbb{P}\left(\bigcap_{k=n}^N A_k^c\right) &= \prod_{k=n}^N (1 - \mathbb{P}(A_k)) \\ &\leq \prod_{k=n}^N e^{-\mathbb{P}(A_k)} = \exp\left\{-\sum_{k=n}^N \mathbb{P}(A_k)\right\}.\end{aligned}$$

得

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(\bigcap_{k=n}^N A_k^c\right) = 0,$$

即  $\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) = 1$ , 故

$$\mathbb{P}(\limsup_n A_n) = 1.$$



## 两种收敛性的关系

注. ① 几乎处处收敛等价于, 对任何  $\epsilon > 0$ , 有

$$\mathbb{P}(\limsup_{n \rightarrow \infty} \{|\xi_n - \xi| > \epsilon\}) = 0.$$

### 推论 1.3.1

如果对任何  $\epsilon > 0$ ,

$$\sum_{n \geq 1} \mathbb{P}(|\xi_n - \xi| > \epsilon) < \infty,$$

那么当  $n \rightarrow \infty$  时, 有  $\xi_n \xrightarrow{\text{a.s.}} \xi$ .

② 几乎处处收敛蕴含依概率收敛, 反之不对.

(反例见书中例 1.3.1.)



## Borel 强大数定律

### 定理 1.3.3

如果  $\{\xi_n, n \geq 1\}$  是独立同分布的随机序列且  $\mathbb{E}\xi_1^4 < \infty$ , 那么  $\{\xi_n\}$  满足强大数定律, 即

$$\frac{1}{n} \sum_{i=1}^n \xi_i \xrightarrow{\text{a.s.}} \mathbb{E}[\xi_1].$$



证. 首先易证 Chebyshev 不等式中的平方可以换成任何正数幂, 例如 4, 即

$$\mathbb{P}(|X| > \epsilon) \leq \frac{\mathbb{E}[|X|^4]}{\epsilon^4}.$$

由推论 1.3.1, 只需证明

$$\sum_n \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n \xi_i - \mathbb{E}\xi_1\right| > \epsilon\right) < \infty.$$



## 回忆展开公式

$$\left(\sum_{i=1}^n x_i\right)^k = \sum_{k_1, \dots, k_n \text{ 非负且和为 } k} \frac{k!}{k_1! \cdots k_n!} x_1^{k_1} \cdots x_n^{k_n},$$

然后由 Chebyshev 不等式,

$$\begin{aligned} \sum_n \mathbb{P} \left( \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - \mathbb{E} \xi_1 \right| > \varepsilon \right\} \right) &\leq \sum_n \frac{1}{(\varepsilon n)^4} \mathbb{E} \left[ \left( \sum_{i=1}^n (\xi_i - \mathbb{E} \xi_i) \right)^4 \right] \\ &= \sum_n \frac{1}{(\varepsilon n)^4} \left( \mathbb{E} \sum_{i=1}^n (\xi_i - \mathbb{E} \xi_i)^4 + 6 \sum_{1 \leq i < j \leq n} \mathbb{E} (\xi_i - \mathbb{E} \xi_i)^2 (\xi_j - \mathbb{E} \xi_j)^2 \right) \\ &= \sum_n \frac{1}{(\varepsilon n)^4} (n \mathbb{E} (\xi_1 - \mathbb{E} \xi_1)^4 + 3n(n-1) (D \xi_1)^2) < \infty, \quad \text{得证.} \end{aligned}$$



### 例 1.3.2 用一笔钱进行投资,

【假设 1】每期银行利率  $r > 0$ ;

① 每期的投资收益率为随机变量  $X > -1$ ;

②  $S_0$  表示初始投资额, 是常数;  $S_n$  表示  $n$  期投资之后的财富.

则

$$S_n = S_0(1 + X_1)(1 + X_2) \cdots (1 + X_n),$$

其中  $X_1, X_2, \dots, X_n$  为第  $1, 2, \dots, n$  期的投资收益率.

【假设 2】投资项目稳定:  $\{X_n\}$  独立且与  $X$  同分布.

问题: 是否存在投资陷阱?

○ 陷阱是指期望很乐观的投资却几乎总是亏损的, 即

对于几乎所有  $\omega \in \Omega$ ,  $S_n(\omega) \rightarrow 0$ .



投资陷阱的存在性: 存在这样的随机收益率  $X$ , 使得

$$(1) \mathbb{E}[S_n] \longrightarrow +\infty;$$

$$(2) S_n \xrightarrow{\text{a.s.}} 0.$$

证. (1) 由期望的性质

$$\mathbb{E}[S_n] = S_0(\mathbb{E}[1 + X])^n = S_0(1 + \mathbb{E}[X])^n,$$

因为  $\mathbb{E}[X] > r > 0$ , 所以

期望收益  $\mathbb{E}[S_n]$  趋于无穷.

$$(2) \text{ 因为 } S_n = S_0 \exp \left( \sum_{i=1}^n \log(1 + X_i) \right). \text{ 由强大数定律,}$$

当  $\mathbb{E}[\log(1 + X)] < 0$  时,

$$\sum_{i=1}^n \log(1 + X_i) \rightarrow -\infty, \text{ 这时 } S_n \xrightarrow{\text{a.s.}} 0. \quad \square$$



## 陷阱条件

$$\begin{cases} \mathbb{E}[1 + X] > 1 \text{ (等价于 } \log \mathbb{E}[1 + X] > 0), \\ \mathbb{E}[\log(1 + X)] < 0. \end{cases}$$

注意,

- ① 很大程度是因为**对数函数的凹性**. 根据 Jensen 不等式,

$$\mathbb{E}[\log(1 + X)] \leq \log(1 + \mathbb{E}[X]).$$

- ② 适当选取  $X$  的分布, 右边是正的, 但左边可能会负的.



假设: 投资有亏损的可能, 即

$$X: \begin{cases} \text{可能大于 } 0 & (\text{盈利}), \\ \text{也可能小于 } 0 & (\text{亏损}). \end{cases}$$

考虑最简单的

$$X \sim \begin{pmatrix} a & b \\ 1-p & p \end{pmatrix}, \text{ 其中 } -1 < a < 0, b > 0,$$

则

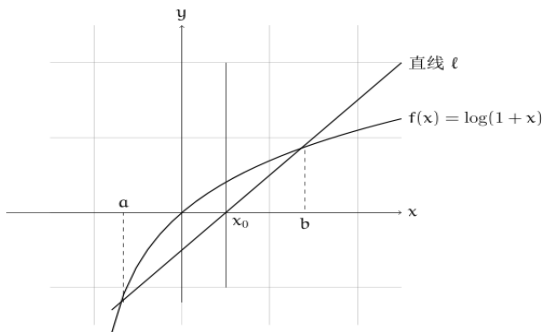
$$\log(1+X) \sim \begin{pmatrix} \log(1+a) & \log(1+b) \\ 1-p & p \end{pmatrix}.$$

此时, 陷阱条件等价于

$$\begin{cases} pb + (1-p)a > 0, \\ p \log(1+b) + (1-p) \log(1+a) < 0. \end{cases}$$



关注函数  $f(x) = \log(1+x)$ :



▷ 对  $p \in (0, 1)$ ,  $\mathbb{E}[X] = pb + (1-p)a \in (a, b)$ , 在直线  $\ell$  上  
对应的点恰是  $\mathbb{E}[\log(1+X)]$ .



## 陷阱区

在  $x > 0$  时,  $f(x) > 0$ , 但直线上有一段却小于 0.

step 1 直线  $\ell$  与  $y = 0$  的交点横坐标

$$x_0 = b - \frac{(b-a)\log(1+b)}{\log(1+b) - \log(1+a)} \in (0, b).$$

step 2 找到  $p$ , 使得  $pb + (1-p)a \in (0, x_0)$ .  
事实上, 解不等式

$$0 < pb + (1-p)a < x_0,$$

得

$$0 < \frac{-a}{b-a} < p < \frac{x_0 - a}{b-a} < 1.$$



## 总结

---

当  $0 < \frac{-a}{b-a} < p < \frac{x_0 - a}{b-a} < 1$  时, 陷阱条件满足:

$$\mathbb{E}[1 + X] > 1 \text{ 但 } \mathbb{E}[\log(1 + X)] < 0,$$

此时出现投资陷阱现象.

#



## 极限与期望交换

问题: 当  $X_n$  几乎处处收敛于某个随机变量时, 是否有

$$\lim_n \mathbb{E}[X_n] = \mathbb{E}[\lim_n X_n]?$$

- 如果能, 我们就说极限与期望可以交换.



## 什么条件下极限和期望可以交换?

- ① 单调收敛定理: 如果  $X_n$  是非负随机序列且单调递增, 则可以交换

$$\mathbb{E}[\lim_n X_n] = \lim_n \mathbb{E}[X_n];$$

- ② Fatou 引理: 如果  $X_n$  是非负随机序列, 则有不等式

$$\mathbb{E}[\liminf_n X_n] \leq \liminf_n \mathbb{E}[X_n];$$

- ③ Lebesgue 控制收敛定理: 如果  $X_n$  几乎处处收敛且存在可积的  $Y$ , 使得  $|X_n| \leq Y$ , 则可以交换

$$\mathbb{E}[\lim_n X_n] = \lim_n \mathbb{E}[X_n].$$



例. 设  $\Omega = (0, 1)$ ,  $\mathcal{F}$  是其中开集生成的  $\sigma$ -代数,  $\mathbb{P}$  为其上 Lebesgue 测度. 令

$$\begin{aligned} X_n(\omega) &= (-1)^n n 1_{(0, \frac{1}{n})}(\omega) \\ &= \begin{cases} (-1)^n n, & \text{当 } \omega \in (0, \frac{1}{n}) \text{ 时,} \\ 0, & \text{当 } \omega \in [\frac{1}{n}, 1) \text{ 时,} \end{cases} \end{aligned}$$

则对任意  $\omega \in \Omega$ ,

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) = 0,$$

$\mathbb{E}[X] = 0$ , 而  $\mathbb{E}[X_n] = (-1)^n$  的极限不存在, 即

$\mathbb{E}[X_n]$  并不收敛于  $\mathbb{E}X$ .



## 矩母函数

: 随机变量函数的数学期望.

定义 1.6.1:

对于随机变量  $X \sim F$ , 若下面的数学期望存在, 则称

$$\psi_X(t) := \mathbb{E}[e^{tX}] = \int_{\mathbf{R}} e^{tx} dF(x), \quad t \in \mathbf{R}.$$

为  $X$  的矩母函数.

常见 分布的矩母函数的计算, 见书中例 1.6.1 (1), (2).



## 随机变量函数的期望

例 1.6.1 (3) 设  $X \sim N(\mu, \sigma^2)$ , 求其矩母函数.

回顾例 1.1.13: 设  $\xi \sim N(0, 1)$ ,  $a \in \mathbf{R}$ , 则有

$$\begin{aligned}\mathbb{E}e^{a\xi} &= \int_{\mathbf{R}} e^{ax} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{-\frac{1}{2}(x^2-2ax)} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{-\frac{1}{2}(x-a)^2} e^{\frac{1}{2}a^2} dx = e^{\frac{1}{2}a^2}.\end{aligned}$$

解. 取  $\xi = \frac{X-\mu}{\sigma}$ , 则由

$$\mathbb{E}[e^{a\frac{X-\mu}{\sigma}}] = \mathbb{E}e^{a\xi} = e^{\frac{1}{2}a^2},$$

再记  $\frac{a}{\sigma}$  为  $a$ , 可得

$$\mathbb{E}e^{aX} = \exp\{\mu a + \frac{1}{2}\sigma^2 a^2\}.$$



## 矩母函数的意义

注. 矩母函数如果存在, 则唯一确定一个概率分布(见书中例 1.6.2).

(1) 通过  $\psi_X(t)$  可求出  $X$  的各阶矩:

$$\mathbb{E}[X^n] = \psi_X^{(n)}(0), \quad n \geq 1.$$

(2) 若  $X, Y$  相互独立, 则

$$\psi_{X+Y}(t) = \psi_X(t) \cdot \psi_Y(t).$$

(见书中例 1.6.3.)



## 母函数

首先 针对实数列  $a = \{a_0, a_1, \dots, a_n, \dots\}$ , 如果幂级数

$$G_a(z) := a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n + \dots,$$

在 0 点的一个非空邻域上收敛, 则称它为数列  $a$  的母函数.

### 定义 1.6.2

设  $\zeta$  是  $\mathbf{Z}_+$ -值随机变量, 则其概率分布律是一个有界数列, 定义

$$G_\zeta(z) := \sum_{n \geq 0} z^n \mathbb{P}(\zeta = n),$$

称之为  $\zeta$  的母函数.



## 母函数的性质

(1)

$$G_{\xi}(z) = \mathbb{E}[z^{\xi}], \quad \lim_{z \uparrow 1} G_{\xi}(z) = \mathbb{P}(\xi < +\infty).$$

(2) 如果  $\xi, \eta$ : 非负整数值随机变量且  $G_{\xi} = G_{\eta}$ , 那么

$\xi$  与  $\eta$  同分布.

(3) 设  $\mathbf{Z}_+$ -值随机变量  $\xi$  的母函数是  $G_{\xi}$ , 那么

$$\mathbb{E}\xi = G'_{\xi}(1), \quad \mathbb{E}\xi^2 = G''_{\xi}(1) + G'_{\xi}(1), \dots$$

事实上,

$$G'_{\xi}(z) = \sum_{n \geq 0} n \mathbb{P}(\xi = n) z^{n-1},$$

$$G''_{\xi}(z) = \sum_{n \geq 1} n(n-1) \mathbb{P}(\xi = n) z^{n-2}.$$



## 母函数的意义

注. 1. 数列  $\{u_n\}$  与  $\{v_n\}$  的卷积, 可定义为如下数列

$$\{u_0 v_n + u_1 v_{n-1} + \cdots + u_n v_0 : n \geq 0\}.$$

2. 如果  $\mathbf{Z}_+$ -值随机变量  $\xi, \eta$  独立, 那么

$\xi + \eta$  的分布列是  $\xi$  与  $\eta$  的分布律的卷积.

### 定理 1.6.1:

独立  $\mathbf{Z}_+$ -值的随机变量  $\xi, \eta$  的和  $\xi + \eta$  的母函数

$$G_{\xi+\eta}(z) = G_{\xi}(z) G_{\eta}(z).$$

(该性质与特征函数, 矩母函数类似.)

