

概率论一点发展历程的介绍

毛翰翔

210110531

2022 年 11 月 16 日

摘要

本文主要介绍了概率的一点发展历程, 主要介绍的有十六十七世纪的古典概型以及中心极限定理大数定理的发展, 以及二十世纪初的概率论的公理化. 十五世纪到十七世纪, 概率论有着显著理论进展, 大数定理和中心极限定理被发现, *Jacob Bernoulli* 和 *Abraham de Moivre* *Pierre Simon Laplace* 这个时期为概率论的应用和理论奠基做出极大的贡献. , 一直到二十世纪初的时候, 概率论又引来显著的发展, 其中俄国数学家 *P. L. Chebyshev* 以及其学生 *A.A. Markov*, 和 *A.N. Kolmogorov* 有着深远的影响. *Markov* 还是随机过程理论的开创者. 本文主要是对我们课内中稍有提到的内容进行介绍, 诸如 Bayes 方法, 随机过程等理论, 恕能力有限, 没有介绍.

1 概率的古代的起源

在古代, 柏拉图和他的学生亚里士多德曾经哲学地讨论 ‘机会’ 这个词. 在公元前 324 年, 一个希腊人, 名曰 *Antimenos*, 第一次创建了保险制度, 能够对一些事件的得失保证一定金额. 我们的日常生活中的各个方面, 比如说健康, 天气, 死亡, 博弈, 都有着 ‘机会’, ‘随机’ 这种概念. 以及物理上的任何测量都带有误差这种本质特性. 概率这个概念起源于游戏的赌博之中. 最早的一个代表人物是著名的物理学家, 数学家: *Gerolamo Cardano* (1501-1576), 后来他成为了位于意大利的 *Bologna* 大学的数学学科的教授.

在十五世纪, 在意大利开始了对于掷色子赌博问题的讨论, 当时有很多对于这样概率问题的文献, 但是都没有说明如何计算概率. *Cardano* 写了一本小书, 名曰 *Liber de Ludo Aleae*. 其中包含了对于概率均值或者说是数学期望的讨论, 并且包含了一个简单的大数定理. 但是, *Cardano* 的这本小书并没有引起多大的重视, 并且没有对概率论的发展做出实质的贡献. 他的小书大约在一个世纪之后 (1633) 才被发行. 他将概率视为一个 0 到 1 之间的数, 并且说明, 如果一个事件的概率为 p , 那么对于一个足够大的数 n , 事件的发生次数将会向 np 靠近.

2 十六十七世纪的概率论的发展

在十六十七世纪, 概率得到了很多的关注, 并且在这个时期, 有很多我们耳熟能详的人物对概率论做出了贡献, 并且在我们的教科书上留下名字. 意大利的著名物理学家天文学家 *Galileo Galilei*, 家喻户晓的伽利略, 提出一种对于掷色子的解法. 这个解法可以说是数学历史上第一个记录在案的概率论成果. 随后 *Baise Pascal* (1623-1662) 和 *Pierre de Fermat* (1601-1665) 两人展开了合作, 他们引入了概率论中重要的概念, 包括: 均值, 条件概率. 并因此, 他们的成果, 可以标志着古典概率论的诞生. 在 *Pascal* 和 *Fermat* 的研究成果上, *Christian Huygens* (1629-1695) 出版了论文 *De Rationiis in Ludo Aleae*, 其中 *Huygens* 介绍了经典的掷色子的问题的解法, 以及一

些多人游戏的概率问题, 补充了 Pascal 和 Fermat 的成果; 并且他还介绍了一个重要的定义, 他将概率定义为某些事件的个数除以全体事件. 并且在 Cardano 的基础之上, 加强了对于随机变量的均值的探索. 随后是 *Jacob Bernoulli* (1654-1704) 其死后 8 年, 他的重要作品 *Ars Conjectandi* 出版.

Bernoulli 提出了概率论中基本的一个定理, 称为 Bernoulli 大数定理. Bernoulli 大数定理非常重要, 具有非常的奠基地位, 以至于后续的概率论发展几乎不能离开 Bernoulli 的大数定理. Bernoulli 的书详细介绍了概率论的应用并且还附带了 Huygens 的题解, 以及 Bernoulli 本人对 Huygens 的论文的思考. 特别地, Bernoulli 关注到了一类特定的重复实验, viz. 我们熟知的 Bernoulli 实验. 我们这里不妨在介绍一下: Bernoulli 实验是 n 次重复独立的, 成功率为 p 二项实验中, 成功次数达到 r 的事件. 我们也能写出这事件的概率:

$$B(r; n, p) = \binom{n}{r} p^r q^{n-r}$$

接下来, 将 r 写为 $np + x$, 并且将 $n - r$ 写为 $nq - x$. 虽然说 n 趋于无穷, 但是我们可以将 r 也定为无穷, 这样的话, x 就可以很小. 使用 *James Stirling* (1693-1770) 的阶乘逼近公式:

$$n! \sim \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} \left(1 + \frac{1}{12n}\right)$$

对于 Bernoulli 分布, 带入 $\binom{n}{r}$ 中的阶乘就有:

$$\begin{aligned} \log B &= \log n! + (np + x) \log p + (nq - x) \log q - \log (np + x)! - \log (nq - x)! \\ &= -\frac{1}{2} \log (2\pi nq) - \frac{1}{2n} \left(\frac{x^2 + x(1-2p)}{pq} \right) \end{aligned}$$

将对数去掉之后, 就有

$$B \approx \frac{1}{\sqrt{2\pi npq}} \exp \left[-\frac{1}{2n} \left\{ \frac{x^2 + x(1-2p)}{pq} \right\} \right]$$

如果说 $|x| \gg |1-2p|$ 就有

$$B(x) = \frac{1}{\sqrt{2\pi npq}} \exp \left(-\frac{x^2}{2npq} \right) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left(-\frac{x^2}{2\sigma^2} \right)$$

我们这样推导出来的就是正态分布 (又称高斯分布), 这样的分布是法国数学家 *Abraham de Moivre* 首先得出的. 甚至先于 Bernoulli 的书, 在 1711 年, de Moivre 出版了 *The Doctrine of Chances*, 虽然说 de Moivre 书中的所画出的直方图已经非常接近钟形, 即正态分布的密度函数, 但是他并没有将这个等式写下来, 也即, 并没有将这个重要的定理: “中心极限定理” 写出来.

后来在 1738 年 *The Doctrine of Chances* 第二版中, 加入了这个定理. 值得一提的是, de Moivre 此时并不了解 Bernoulli 的工作, 书中并没有直接提到 Bernoulli 分布, 但其分析的确是 Bernoulli 分布.

3 概率论的公理化

对于大数定理和中心极限定理, 俄国的 *P. L. Chebyshev* (1821-1894) *A.A. Markov* (1856-1922) 做出了极大的贡献, Chebyshev 的矩法极大的简化了部分大数定理的证明. 而 Markov 将中心极限定理的条件极大地放宽. 但是这里要讲的是一个, 二十世纪概率论发展中最有影响力的人物, *A.N. Kolmogorov* 他对众多领域做出了非常杰出的贡献: 傅里叶级数和三角级数, 测度论,

集合论, 积分, 数理逻辑, 逼近理论, 概率论, 拓扑学, 随机过程, 统计, 信息论, 动力系统, 自动机理论, 微分方程, 弹道分析, Hilbert 的 13 大问题.

在 1933 年德国, Kolmogorov 的 *Foundations of Probability Theory* 一书出版, 首先建立了概率论的现代公理基础. 我们这里将一点公理简要说明一下:

定义 1. 概率空间是一个三元组 $(\Omega, \mathcal{F}, \mathcal{P})$, 其中 Ω 是样本空间, 其中元素 ω 被称为是基本事件, \mathcal{F} 是 *sigma-field*, 他是 Ω 的一个子集族, 并且满足以下性质:

1. $\forall A \in \mathcal{F} \implies A^c \in \mathcal{F}$
2. $\forall A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$
3. $\forall A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$
4. $\Omega \in \mathcal{F}$

不难看出, 如果条件 1 满足, 条件 2 和 3 实际上是等价的. 这样的 \mathcal{F} 也称为 sigma algebra.

定义 2. 概率空间的 \mathcal{P} 满足:

1. $\mathcal{P} : \mathcal{F} \rightarrow [0, 1]$
2. $\mathcal{P}(\Omega) = 1$
3. 对于不相交的 $E_n \in \mathcal{F}$ 有 $\mathcal{P}(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mathcal{P}(E_n)$

条件 3 称为可数可加性. 并且 \mathcal{P} 称为是概率测度.

并且在这个基础之上我们还能够精确地定义随机变量:

定义 3. 随机变量 X 是满足以下条件的实函数:

1. $X : \Omega \rightarrow \mathbb{R}$
2. $\forall A \in \mathcal{B}, X^{-1}(A) \in \mathcal{F}$

其中 \mathcal{B} 是 \mathbb{R} 上的 *Borel* 集

定义 4. \mathbb{R} 上的 *Borel* 集定义为

$$\mathcal{B} = \bigcap_{\alpha \in \Lambda} B_{\alpha}$$

其中 B_{α} 是包含了 \mathbb{R} 上所有开集的 *sigma field*.

于是说我们常常书写的符号 $P(X = a)$ 或者是 $P(X \in A)$, 其意思是

$$P(X \in A) = \mathcal{P}(X^{-1}(A)), A \in \mathcal{B}$$

我们能够知道概率空间的概念就是测度空间的一种, 在这种意义之下, 期望以及矩是通过 Lebesgue 积分计算的. 此后, 概率论的基础就变为了实分析和测度论.