概率论与数理统计

金玲飞

复旦大学计算机学院
Email: lfjin@fudan.edu.cn

2018.9.11

- 预备知识：集合论，排列组合，微积分
- 成绩评定：平时40%（作业，期中等），期末60%.
- 教材:《概率论与数理统计》第二版，齐民友主编，高等教育出版社。

李宗义
15302010004@fudan.edu.cn

陆怡
15302010044@fudan.edu.cn

- 第一部分：概率论 (probability theory)（10周）
  —-研究随机现象的规律性

- 第二部分：数理统计 (mathematical statistics)（5周）
  —研究随机现象的数据收集与处理（怎样确定一个随机变量的分布及其参数）

概率论背景介绍

- 考古上，最早的骰子是出现在埃及。在两千多年前古埃及的骰子被称为astragal。
- 考古学家曾在出土的古埃及坟墓的壁上，绘有以羊的后足跟制成的称为astragal之赌具的赌戏。这种骨头有四个面，并不对称，每次投掷会落在四个面之一方。
- 在我国，山东青州战国齐墓出土的是最早的。
- 真正中国本土国产骰子14面和18面，秦皇陵出土的骰子为中国正宗本土骰子，上刻汉字，秦汉以来的多面骰子随着文化交流，到后来点数一说传入中国，接着中西结合，才有了现在我们常见的骰子。

概率论背景介绍

- 起源于17世纪。一般把1654年作为概率论诞生的一年。法国的数学家帕斯卡Pascal和费马Fermat就机会博弈中的问题进行了通信讨论。后来惠更斯也加入。

布莱士·帕斯卡（Blaise Pascal ）：法国数学家，物理学家

皮耶·德·费马（Pierre de Fermat）：法国律师，业余数学家

概率论的发展史

- 1657年，惠更斯Huygens发表了《论赌博中的计算》，这本书成为了最早的概率论著作.

——-这一时期称为组合概率时期，研究古典概率。

- 1713年，瑞士数学家伯努利（J. Bernoulli）著作《猜度术》出版。他证明了大数定律，作为大数定律的最早形式而在概率论发展史上占有重要地位。

概率论的发展史

- 19世纪，法国数学家拉普拉斯将古典概率论向近代概率论进行推进，引入了更有力的数学分析工具，将概率论推向一个新的发展阶段。
- 19世纪后期，极限理论的发展成为概率论研究的中心课题，1866年，俄国数学家Chebyshev建立大数定律，使伯努利定理成为其特例。

概率论的发展史

- 19世纪，法国数学家拉普拉斯将古典概率论向近代概率论进行推进，引入了更有力的数学分析工具，将概率论推向一个新的发展阶段。
- 19世纪后期，极限理论的发展成为概率论研究的中心课题，1866年，俄国数学家Chebyshev建立大数定律，使伯努利定理成为其特例。
- 20世纪初完成的勒贝格测度与积分理论及随后发展的抽象测度和积分理论，为概率公理体系的建立奠定了基础。
- 1933年，Kolmogorov在《概率论基础》一书中首次给出了概率的测度论定义和一套严密的公理体系。他的公理化方法成为现代概率论的基础，使概率论成为严谨的数学分支。

概率统计几乎遍及所有科学技术领域。

原因	天数
饮酒	-130
未婚男性	-3500
未婚女性	-1600
惯用左手	-3285
30%超重	-1300
家有烟雾警报	+10

（1）假设生男生女概率相等。在一个重男轻女的国家里，每个家庭都想生男孩，如果他们生的孩子是女孩，就再生一个，直到生下的是男孩为止。这样的国家，男女比例会是多少？

(2)甲乙两人进行乒乓球比赛，每局甲胜的概率为0.4。假设各局胜负是独立的，对甲而言，采用三局两胜还是五局三胜比较有利。

(2)甲乙两人进行乒乓球比赛，每局甲胜的概率为0.4。假设各局胜负是独立的，对甲而言，采用三局两胜还是五局三胜比较有利。

当甲处于劣势的时候，比赛局数越多越不利！

第一章 随机事件与概率

自然界和人类社会中的两种现象：

- **确定性现象**：在一定条件下必然发生的结果
- **随机现象**：在实验前不能预知明确的结果，可能出现这样或那样的结果。

自然界和人类社会中的两种现象：

- **确定性现象：** 在一定条件下必然发生的结果
- **随机现象：** 在实验前不能预知明确的结果，可能出现这样或那样的结果。

概率论与数理统计——研究随机现象的统计规律

随机试验

随机试验的三个特征：

- 可在相同的条件下重复进行；
- 试验的所有可能的结果不止一个，并且是事先知道的；
- 试验之前无法知道哪一个结果会出现。

随机试验

随机试验的三个特征：

- 可在相同的条件下重复进行；
- 试验的所有可能的结果不止一个，并且是事先知道的；
- 试验之前无法知道哪一个结果会出现。

一些记号：

- 随机试验，$\mathcal{E}$
- 样本点（基本事件），ω：试验的每一个可能结果
- 样本空间，Ω：试验的所有可能的结果

随机实验的例子：

研究随机现象首先要列出样本空间

- $\mathcal{E}_1$: 抛一枚硬币，观察正反面出现的情况。
- $\mathcal{E}_2$: 将一枚硬币抛三次，观察正反面出现的情况。

随机实验的例子：

研究随机现象首先要列出样本空间

- $\mathcal{E}_1$: 抛一枚硬币，观察正反面出现的情况。
- $\mathcal{E}_2$: 将一枚硬币抛三次，观察正反面出现的情况。

- $\mathcal{E}_3$:观察某一城市某一天共发生的车祸的次数。
- $\mathcal{E}_4$:任抽一台电视，测它的寿命。

随机事件

考虑某一特定条件的样本点是否出现。

定义 (随机事件 random event)

把样本空间的满足某一特定条件的样本点或是具有某种特征的样本点组成的子集称为随机事件（事件）。记为 $A, B, C, \ldots$ 。

随机事件

考虑某一特定条件的样本点是否出现。

定义 (随机事件 random event)

把样本空间的满足某一特定条件的样本点或是具有某种特征的样本点组成的子集称为随机事件（事件）。记为 $A, B, C, \ldots$。

- 随机事件发生：组成随机事件的一个样本点发生
- 必然事件，全集
- 不可能事件，空集

事件的关系和运算

事件的关系：

- 事件的包含：$A \subset B$ 或者$B \supset A \Leftrightarrow$。 若事件A发生，则事件B发生。
- 事件的相等：若$A \subset B$且$A \supset B$。

事件的关系和运算

事件的关系：

- 事件的包含：$A \subset B$ 或者$B \supset A \Leftrightarrow$。 若事件A发生，则事件B发生。
- 事件的相等：若$A \subset B$且$A \supset B$。
- 逆事件或对立事件：$\overline{A} = \Omega - A$ 表示事件A不发生。
- 不相容: $AB = \varnothing$，即 A与B不可能同时发生。

事件的关系和运算

事件的关系：

- 事件的包含：$A \subset B$ 或者$B \supset A \Leftrightarrow$。 若事件A发生，则事件B发生。
- 事件的相等：若$A \subset B$且$A \supset B$。
- 逆事件或对立事件：$\overline{A} = \Omega - A$ 表示事件A不发生。
- 不相容: $AB = \varnothing$，即 A与B不可能同时发生。

注：也可用文氏图表示。

事件的运算：

- 事件的并（加）： $A \cup B$或$A+B$
- 事件的交（积）： $A \cap B$或AB
- 事件的差： $A-B$

事件的运算满足和集合的运算一样的运算规律:

- 交换律: $A \cup B = B \cup A$, $AB = BA$
- 结合律:

$$A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C$$

- 分配律:

$$A{\cup}(B{\cap}C) = (A{\cup}B){\cap}(A{\cup}C), A{\cup}(B{\cap}C) = (A{\cup}B){\cap}(A{\cup}C)$$

事件的运算满足和集合的运算一样的运算规律:

- 交换律: $A \cup B = B \cup A$, $AB = BA$
- 结合律:

$$A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C$$

- 分配律:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

- 德摩根律:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{AB} = \overline{A} \cup \overline{B}$$

常用的性质：

不相容分解

$(1) A \cup B = A \cup \overline{A}B$
$(2) A = AB \cup A\overline{B}$

常用的性质：

不相容分解

$$(1) A \cup B = A \cup \overline{A}B$$
$$(2) A = AB \cup A\overline{B}$$

其它性质

$$(3) A - B = A\overline{B} = A - AB$$
$$(4) \overline{\overline{A}} = A$$

例：试用A, B, C表示下列事件

(1) A, B 发生,C不发生；

(2) 都不发生；

(3) 恰有一个发生；

(4) 至少两个发生；

(5) 至多有一个发生；

(6) 不全发生；

事件域和事件的数学定义

定义 (1.1.1 事件域)

设Ω为样本空间，$\mathcal{F}$为Ω的某些子集组成的集合，如果它满足：

- $\Omega \in \mathcal{F}$;
- $A \in \mathcal{F} \Rightarrow \overline{A} \in \mathcal{F}$;
- $A_n \in \mathcal{F}, n \geq 1 \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

则称$\mathcal{F}$为（Ω上的）事件域（或 σ*-field*），称$\mathcal{F}$中的集合为事件。

事件域和事件的数学定义

定义 (1.1.1 事件域)

设Ω为样本空间，$\mathcal{F}$为Ω的某些子集组成的集合，如果它满足：

- $\Omega \in \mathcal{F}$;
- $A \in \mathcal{F} \Rightarrow \overline{A} \in \mathcal{F}$;
- $A_n \in \mathcal{F}, n \geq 1 \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

则称$\mathcal{F}$为（Ω上的）事件域（或 σ-*field*），称$\mathcal{F}$中的集合为事件。

- 最小事件域：
- 最大事件域：

1.2 概率与频率

确定概率的方法

- 频率的方法
- 古典的方法
- 几何的方法
- 主观的方法

概率的公理化定义

定义 (概率)

设Ω为样本空间，$\mathcal{F}$为Ω上的事件域，P为定义在$\mathcal{F}$上的实值函数，如果P满足，

- 非负性: $\forall A \in \mathcal{F}, P(A) \geq 0$
- 规范性: $P(\Omega) = 1$
- 可列可加性：对$\mathcal{F}$ 中任何可列个无穷多个互不相容的事件$A_1, \ldots, A_k, \ldots$

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

则称P为$\mathcal{F}$上的概率测度，简称为概率；称$(\Omega, \mathcal{F}, P)$为概率空间；称$P(A)$为事件A的概率。

例子 (1.2.1)

考虑掷骰子试验，样本空间为$\Omega = \{1, 2, 3, 4, 5, 6\}$，事件域$\mathscr{F}$取为$\Omega$的所有子集全体。若骰子是均匀的，令$P(A) = \frac{A\text{中元素的个数}}{6}$，$A \in \mathscr{F}$。则$(\Omega, \mathscr{F}, P)$为概率空间。

例子 (1.2.1)

考虑掷骰子试验，样本空间为$\Omega = \{1,2,3,4,5,6\}$，事件域$\mathscr{F}$取为Ω的所有子集全体。若骰子是均匀的，令$P(A) = \frac{A\text{中元素的个数}}{6}$，$A \in \mathscr{F}$。则$(\Omega, \mathscr{F}, P)$为概率空间。

例子 (1.2.2)

设样本空间$\Omega = [0,1]$，事件域$\mathscr{F}$由$[0,1]$中所有可求长的子集构成。令$P(A)$等于A的长度，$A \in \mathscr{F}$，则$(\Omega, \mathscr{F}, P)$为概率空间。
类似地，$\Omega = [0,1] \times [0,1]$为平面中单位正方形，$\mathscr{F}$为$[0,1] \times [0,1]$中具有面积的子集全体，$P(A) = A$的面积，则$(\Omega, \mathscr{F}, P)$为概率空间。

概率的基本性质

设$(\Omega, \mathcal{F}, P)$为概率空间，概率P具有如下性质：

- **不可能事件的概率为0：** $P(\phi) = 0$；
- **有限可加性：** 对$\mathcal{F}$中任何有限多个互不相容的事件$A_1, \ldots, A_n$，

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

概率的基本性质

设$(\Omega, \mathcal{F}, P)$为概率空间，概率P具有如下性质：

- **不可能事件的概率为0：** $P(\phi) = 0$；
- **有限可加性：** 对$\mathcal{F}$中任何有限多个互不相容的事件$A_1, \ldots, A_n$，

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

- **可减性：** 若$A \subset B$,则$P(B - A) = P(B) - P(A)$;

概率的基本性质

设$(\Omega, \mathcal{F}, P)$为概率空间，概率P具有如下性质：

- **不可能事件的概率为0：** $P(\phi) = 0$；
- **有限可加性：** 对$\mathcal{F}$中任何有限多个互不相容的事件$A_1, \ldots, A_n$，

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

- **可减性：** 若$A \subset B$,则$P(B - A) = P(B) - P(A)$;
- **单调性：** 若$A \subset B$,则$P(A) \leq P(B)$;
- **逆事件概率公式：** $P(\overline{A}) = 1 - P(A)$。

加法公式

- 若$A, B \in \mathcal{F}$，则

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

加法公式

- 若$A, B \in \mathcal{F}$，则

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

- 若$A_1, \ldots, A_n \in \mathcal{F}$，则

$$p\left(\cup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k) - \sum_{1 \leq i < j \leq n} P(A_i A_j)$$

$$+ \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) - \ldots + (-1)^{n-1} P(A_1 A_2 \cdots A_n)$$

例子 (1.2.4)

设$p = P(A), q = P(B), r = P(A \cup B)$。
求$P(A\overline{B}), P(\overline{A} \cap \overline{B})$。

概率的连续性

定义

若事件序列$\{A_n : n \geq 1\}$满足$A_1 \subset A_2 \subset \cdots$，则称该事件序列是单调增的，称$\bigcup_{n=1}^{\infty} A_n$为单调增序列$\{A_n : n \geq 1\}$的极限；若事件序列$\{A_n : n \geq 1\}$满足$A_1 \supset A_2 \supset \cdots$，则称该事件序列是单调降的，称$\bigcap_{n=1}^{\infty} A_n$为单调降序列$\{A_n : n \geq 1\}$的极限。

概率的连续性

定义

若事件序列$\{A_n : n \geq 1\}$满足$A_1 \subset A_2 \subset \cdots$，则称该事件序列是单调增的，称$\bigcup_{n=1}^{\infty} A_n$为单调增序列$\{A_n : n \geq 1\}$的极限；若事件序列$\{A_n : n \geq 1\}$满足$A_1 \supset A_2 \supset \cdots$，则称该事件序列是单调降的，称$\bigcap_{n=1}^{\infty} A_n$为单调降序列$\{A_n : n \geq 1\}$的极限。

- 不论$\{A_n : n \geq 1\}$是单调增还是单调降，A_n的概率均收敛到它极限的概率，也称上，下连续性。

概率的连续性

定理

设$(\Omega, \mathcal{F}, P)$为概率空间，概率P具有如下性质：

(1) 下连续性：若$\{A_n : n \geq 1\}$是单调增的事件序列，则

$$\lim_{n\to\infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n)$$

(2) 上连续性：若$\{A_n : n \geq 1\}$是单调降的事件序列，则

$$\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$$

概率的连续性

定理

设$(\Omega, \mathcal{F}, P)$为概率空间，概率P具有如下性质：

(1) 下连续性：若$\{A_n : n \geq 1\}$ 是单调增的事件序列，则

$$\lim_{n\to\infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n)$$

(2) 上连续性：若$\{A_n : n \geq 1\}$ 是单调降的事件序列，则

$$\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$$

由可列可加性可以推出有限可加性和下连续性。

确定概率的频率方法

定义 (1.2.1)

设Ω是随机试验 $\mathcal{E}$的样本空间，事件 $A \subset \Omega$。在相同条件下将试验$\mathcal{E}$重复n 次，以$n(A)$表示事件A在这n次试验中发生的次数，称 $\frac{n(A)}{n}$为A在这n次试验中发生的频率，记为

$$f_n(A) = \frac{n(A)}{n}$$

从频率到概率

- 频率的稳定性: 实验次数$n \to \infty, f_n(A) \to p$
- 频率具有概率的性质
  - 非负性
  - 规范性
  - 可列可加性

从频率到概率

- 频率的稳定性: 实验次数$n \to \infty, f_n(A) \to p$
- 频率具有概率的性质
  - 非负性
  - 规范性
  - 可列可加性

例子： 字母使用率，研究女婴的出生率等（利用频率给出概率近似值）

1.3 古典概型与几何概型

古典概型

历史上最早使用的方法之一。

定义 (古典概型，又称有限等可能概型)

具有如下两个特征的随机试验称为有限等可能概型，

- 试验只有有限个可能的结果；
- 试验的每个基本事件发生的可能性相同。

古典概型

历史上最早使用的方法之一。

定义 (古典概型，又称有限等可能概型)

具有如下两个特征的随机试验称为有限等可能概型，

- 试验只有有限个可能的结果；
- 试验的每个基本事件发生的可能性相同。

对任何事件 $A \subset \Omega$，

$$P(A) = \frac{\text{A中元素的个数}}{\Omega\text{中元素的个数}}$$

方法：采用排列组合等进行计数。