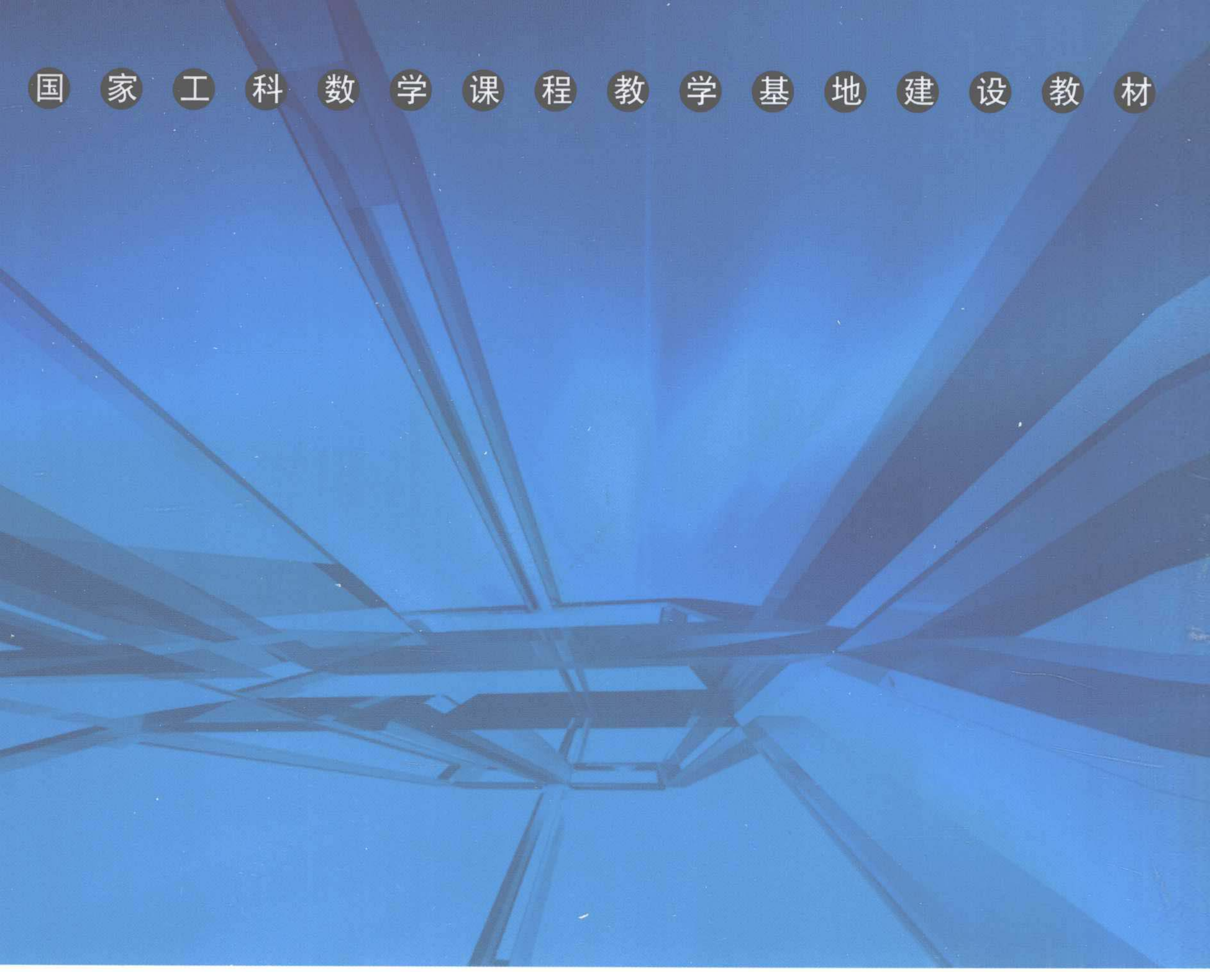
国家工科数学课程教学基地建设教材



概率论与数理统计

◎ 华南理工大学理学院数学系

◎ 何春雄　龙卫江　朱锋峰　编

高等教育出版社

HIGHER EDUCATION PRESS

国家工科数学课程教学基地建设教材

概率论与数理统计

Gailülun yu Shuli Tongji

华南理工大学理学院数学系
何春雄　龙卫江　朱锋峰　编

高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

内容提要

本书是高等学校非数学专业“概率论与数理统计”课程的教材，以微积分学和线性代数为基础，介绍概率论和数理统计的基础知识。概率论部分包括概率的定义、条件概率与独立性、随机变量（包括随机向量）的分布和数字特征、大数定律和中心极限定理。数理统计部分包括抽样分布、参数估计和假设检验及 R 软件简介。书中尽量略去繁复的推导和计算，重点介绍和解释概率论与数理统计的基本概念、基本思想和基本方法，同时将 R 软件的部分内部函数和简单程序用于例题的计算，每章末有小结与注记，各章配有习题并在书末附有习题参考答案。

本书可作为高等学校理工、管理、经济、金融等专业“概率论与数理统计”课程的教材。读者只需具有初等微积分学和线性代数以及简单的计算机知识，就可以阅读和学会使用 R 软件进行数据处理。课堂教学中可不专门讲解 R 软件，教师只需在例题的计算中做简单演示即可。

图书在版编目（CIP）数据

概率论与数理统计／何春雄，龙卫江，朱锋峰编．-- 北京：高等教育出版社，2012. 2

ISBN 978-7-04-034020-4

Ⅰ．①概… Ⅱ．①何… ②龙… ③朱… Ⅲ．①概率论－高等学校－教材②数理统计－高等学校－教材 Ⅳ．①O21

中国版本图书馆 CIP 数据核字（2011）第 273620 号

策划编辑	王　强	责任编辑	杨　波	封面设计	于文燕	版式设计	王艳红
插图绘制	尹文军	责任校对	刘　莉	责任印制	尤　静		

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街4号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	北京宏信印刷厂	网上订购	http://www.landraco.com
开　　本	787mm×960mm　1/16		http://www.landraco.com.cn
印　　张	15.75	版　　次	2012 年2月第 1 版
字　　数	290 千字	印　　次	2012 年2月第1次印刷
购书热线	010-58581118	定　　价	21.80 元

物 料 号　34020-00

前　　言

概率论与数理统计是研究随机现象的统计规律的专门学科。它既有广泛的实际应用，又是数学科学领域内重要的分支之一。长期以来，概率论与数理统计的理论和方法被广泛应用于工业、农业、军事、经济、金融、管理和自然科学研究等诸多领域，特别是现代信息技术的高速发展，使该学科在解决各类实际问题中发挥越来越重要的作用。作为一门独特的数学分支，在其研究不断深入的同时，还与其他数学分支互相渗透，促进了整个数学学科的发展。

概率论与数理统计作为大学理工、管理、经济等专业的数学基础课程，由于学生数学基础和课时的限制，不可能较为深入地介绍该学科的学术成果，只能介绍其基本概念和思想，以期学生能借助对该学科概念、方法及思维方式的初步理解和掌握，去分析和对待无处不在的随机现象。所以，本书的内容只能说是初等的，并且很不全面。

本书的内容分为三大部分，即概率论、数理统计和 R 软件简介。应当说，虽然概率论和数理统计的思想方法是一致的，但分析和处理问题的前提或角度有所不同。概率论是在对随机现象有基本认知的前提下，进行演绎推理；而数理统计则试图通过试验去认知 (通过局部推断整体) 随机现象。另外，数理统计处理问题的思路或想法往往来自直观想法或概率论的有关结果，数理统计方法优劣的评价也依靠概率论来论证，从这一点讲，概率论是数理统计的基础。R 软件则是数理统计的计算工具，它是众多统计软件中小巧而开放的软件，对初学者使用已经足够。按目前学生计算机方面的能力，本部分内容应安排自学，教师只在讲解例题时作简单演示即可。因此我们在数理统计的例题计算中已引用了 R 软件中的诸如 p 值、q 值、d 值以及样本均值、样本方差等内部函数，还附以简单 R 程序，以使读者对该软件有较早接触和体会。如条件允许，读者在做习题时就可以用它，这也是编者所期望的！考虑到学生未必都有使用计算机的条件，在书末附加了数理统计常用统计分布表。

本书写作过程中，理论和方法部分由何春雄执笔，习题及解答由朱锋峰负责完成，R 软件介绍及例题的计算由龙卫江负责实现。限于编者的知识和能力，谬误与不足之处在所难免，恳请读者批评指正。

编　者

2011 年 10 月于华南理工大学理学院

目　　录

第一章　随机事件与概率

§1.1　随机现象与随机试验

自然界和人类社会的世间万象, 大体可分为两类, 一类是确定性现象, 另一类是偶然性现象. 我们所说的确定性现象是指, 当满足一定条件时, 该类现象的结果是可以预见的. 比如, 明天太阳会从东方升起, 掷一枚石子会落地等等. 而偶然性现象则指那种结果无法预见的现象. 比如某强台风未来 24 小时中心最大风速和走向, 明天股市的涨与跌, 掷一颗骰子出现的点数等等. 在偶然性现象中, 有一部分是可以在基本相同的条件下重复观测或重复试验的, 并且随着观测或试验次数的增多, 会看出试验的不同结果出现的可能性有大有小, 也就是我们通常所讲的呈现出统计规律, 这种现象称为**随机现象**. 仔细想来, 我们生活的世界里, 随机现象是大量存在的, 并且影响着我们生活的方方面面. 为了刻画和研究随机现象, 数学家经过不懈的努力, 开创了概率论这门重要而独特的数学分支. 换言之, **概率论是研究随机现象的统计规律的科学**, 它的独特之处在于用确定的数学研究非确定的现象 (随机现象).

1.1.1　随机试验与随机事件

用确定性的数学研究随机现象, 从何入手呢? 人们在建立数学模型时, 引进了三个最基本的概念, 它们是随机试验、随机事件和概率.

我们把对随机现象的一次观测或实际实验, 统称为**随机试验**, 其基本意义是

(1) 试验可以在基本相同的条件下大量重复.

(2) 试验会出现的那些结果是可预知的.

(3) 每次试验将出现哪一个结果是无法预知的.

我们将随机试验的结果称为**随机事件**, 简称为**事件**. 而**概率**则是随机事件发生的可能性大小的一种度量, 这是用概率论学科研究随机现象的关键切入点. 然而如何规定或定义这个度量却非易事, 我们将在后文中仔细讨论.

例 1.1.1　掷一颗骰子, 观察其落定后朝上面出现的点数, 这就是一个随机试验. 而 $A=\{$出现的点数为奇数$\}$, $B=\{$出现的点数为偶数$\}$, $C=\{$出现的点数小于 3$\}$ 等都是事件. 若骰子均匀且每次掷法随意, 则依我们后面介绍的概率的

古典定义, A, B, C 的概率应分别为 $P(A)=\frac{1}{2}, P(B)=\frac{1}{2}, P(C)=\frac{1}{3}$. □

例 1.1.2 设地铁每 5 分钟开出一列, 观测一位乘客的等待时间, 这也是一个随机试验. 而 $A=$\{等待时间不超过 2 分钟\}, $B=$\{等待时间多于 2 分钟\}, $C=$\{等待时间介于 1 分钟到 2 分钟之间\} 等都是事件. 若乘客随意到达, 则依我们后面介绍的概率的几何定义, A, B, C 的概率应分别为 $P(A)=\frac{2}{5}, P(B)=\frac{3}{5}, P(C)=\frac{1}{5}$. □

1.1.2 事件的关系与运算

前已指出, 概率论的研究对象是随机现象的统计规律, 最终回答的问题归结为一个事件发生的可能性的大小, 即事件发生的概率有多大. 在分析计算一个较为复杂的事件的概率时, 人们往往用较为简单事件的运算来表达复杂事件, 进而通过计算简单事件的概率来计算复杂事件的概率. 所以本小节介绍事件之间的关系与运算.

在观察一个随机现象时, 某个结果 (或事件) 每次都会出现 (或发生), 这种事件我们称之为**必然事件**, 而某个结果 (或事件) 每次都不出现, 这种事件我们称之为**不可能事件**. 比如, 在例 1.1.1 中的 \{出现的点数不超过 6\} 和例 1.1.2 中的 \{等待时间不超过 5 分钟\}, 都是必然事件. 而 \{出现的点数为 7\} 和 \{等待时间超过 5 分钟\} 都是不可能事件.

通常用 Ω 表示必然事件, 而用 $\varnothing$ 表示不可能事件. 这两个事件其实不是随机的, 只是为了以后表述的方便而引进的, 这一点, 与微积分学中把常数也叫变量是雷同的.

事件的关系

在观察随机现象时, 一次试验中若事件 A 发生, 则事件 B 也发生, 我们称B **包含** A, 记为 $A\subset B$. 若 $A\subset B$ 且 $B\subset A$, 则称 A 与 B **相等**, 记为 $A=B$. 一次试验中事件 A 不发生这样的结果, 称为 A 的**对立事件**, 记为 $\overline{A}$. 若每次试验中, A 和 B 不会同时出现, 则称 A 与 B**互不相容** .

在例 1.1.1 中, \{出现的点数小于 3\}$\subset$\{出现的点数小于 5\}, $\overline{A}=B$, 且 A 与 B 互不相容. 在例 1.1.2 中, \{等待时间多于 3 分钟\}$\subset$\{等待时间多于 2 分钟\}, $\overline{A}=B$, 且 B 与 C 互不相容. 另外, 规定 $\varnothing\subset\Omega, \overline{\varnothing}=\Omega$.

事件的运算

1. 事件的积 (或交).

设 A 和 B 为任意两事件, 一次试验中, 若 A 和 B 都出现, 这样的结果称为 A 与 B 的**积** (或**交**). 记为 $A\cap B$ (或 AB). 换言之, $A\cap B$ 表示 A 和 B 都出现

这样的试验结果.

例 1.1.3 (例 1.1.1 续) 沿用例 1.1.1 的有关记号, 则有

$A \cap C = \{$出现点数为 1$\}$,

$B \cap C = \{$出现的点数为 2$\}$,

$A \cap B = \varnothing$. □

2. 事件的和 (或并)

设 A 和 B 为任意两事件, 一次试验中, 若 A 或 B 至少出现一个, 这样的结果称为 A 与 B 的**和** (或**并**). 记作 $A \cup B$. 也就是说, $A \cup B$ 表示 A 和 B 至少出现一个这样的结果.

例 1.1.4 (例 1.1.2 续) 沿用例 1.1.2 的有关记号, 则有

$A \cup C = A$,

$B \cup C = \{$等待时间超过 1 分钟$\}$,

$A \cup B = \Omega$. □

3. 事件的差

设 A 和 B 为任意两事件, 一次试验中, 若 A 出现但 B 不出现, 这样的结果称为 A 与 B 的**差**. 记作 $A \backslash B$(或 $A - B$). 也就是说, $A \backslash B$ 表示 A 出现但 B 不出现这样的结果.

例 1.1.5 (例 1.1.1 续) 沿用例 1.1.1 的有关记号, 则有

$A \backslash C = \{$出现的点数为 3,5$\}$,

$C \backslash A = \{$出现的点数为 2$\}$,

$A \backslash B = A$, $B \backslash A = B$. □

例 1.1.6 设有 A, B, C 三个事件, 则如下的事件可以用事件的运算表示为

(1) $\{A$ 和 B 都发生但 C 不发生$\} = AB\overline{C} = AB\backslash C$,

(2) $\{A$ 和 B 都不发生但 C 发生$\} = \overline{A}\,\overline{B}C = C\backslash(A \cup B) = C(\overline{A \cup B})$,

(3) $\{A, B$ 和 C 都发生$\} = ABC$,

(4) $\{A, B$ 和 C 都不发生$\} = \overline{A}\,\overline{B}\,\overline{C}$,

(5) $\{A, B$ 和 C 不都发生$\} = \overline{ABC}$,

(6) $\{A, B$ 和 C 中恰有两个发生$\} = AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$,

(7) $\{A, B$ 和 C 中至少有两个发生$\} = AB \cup BC \cup AC = ABC \cup AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$,

(8) $\{A, B$ 和 C 中最多有一个发生$\} = \overline{AB \cup BC \cup AC} = \overline{A}\,\overline{B}\,\overline{C} \cup \overline{A}\,\overline{B}C \cup \overline{A}B\overline{C} \cup A\overline{B}\,\overline{C}$. □

例 1.1.7 在飞碟射击比赛中, 假设规定每位射手连续射击三次. 试用事件的运算表示下列各事件.

(1) $A=$\{三次射击都射中\}.

(2) $B=$\{三次射击中只射中两次\}.

(3) $C=$\{三次射击中至少有一次未射中\}.

解 记 $A_i=$ \{第 i 次射中目标\}, $i=1,2,3$.

(1) $A=A_1A_2A_3$.

(2) \{三次射击中只射中两次\} 就是三次中有一次未射中, 而其余两次射中, 所以

$$B=\overline{A_1}A_2A_3\cup A_1\overline{A_2}A_3\cup A_1A_2\overline{A_3}.$$

(3) \{三次射击中至少有一次未射中\} 就是 "或者全未射中, 或者恰有一次未射中, 或者恰有两次未射中", 也就是 "不是三次都射中", 亦即 "$\overline{A_1}$, $\overline{A_2}$ 和 $\overline{A_3}$ 中至少有一个发生", 所以

$$\begin{aligned}C=&\left(\overline{A_1}\ \overline{A_2}\ \overline{A_3}\right)\cup\left(\overline{A_1}A_2A_3\cup A_1\overline{A_2}A_3\cup A_1A_2\overline{A_3}\right)\cup\\&\left(\overline{A_1}\ \overline{A_2}A_3\cup\overline{A_1}A_2\overline{A_3}\cup A_1\overline{A_2}\ \overline{A_3}\right)\\=&\overline{A_1A_2A_3}=\overline{A_1}\cup\overline{A_2}\cup\overline{A_3}.\end{aligned}$$

□

例 1.1.8 图 1.1 是一个开关电路, 试用各开关的 "开" 或 "闭" 表示 "灯亮". 记 A_i 为第 i 个开关闭合, $i=1,2,3,4$. $B=$ \{灯亮\}. 由于要使电路接通当且仅当 "开关 1 和 2 同时闭合或者开关 3 或者开关 4 至少有一个闭合", 所以

$$B=A_1A_2\cup A_3\cup A_4.$$

□

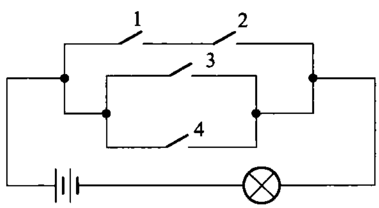


图 1.1 开关电路示意图

事件运算的法则

前面我们定义了事件的积 (交)、和 (并)、差及对立等运算, 与算术运算和极限运算等类似, 为分析计算较为复杂事件的概率, 我们还需熟悉事件运算的法则. 这些运算法则, 读者在例 1.1.6 和例 1.1.7 中已经有所体会了.

首先, 对于积 (交), 和 (并) 运算有

(1) **结合律** 由于 $(A\cup B)\cup C$ 和 $A\cup(B\cup C)$ 都表示 A, B, C 至少发生一个, 所以

$$(A\cup B)\cup C=A\cup(B\cup C). \tag{1.1.1}$$

而 $(A\cap B)\cap C$ 和 $A\cap(B\cap C)$ 都表示 A, B, C 都发生, 所以

$$(A\cap B)\cap C=A\cap(B\cap C). \tag{1.1.2}$$

(2) **交换律**

$$A\cup B=B\cup A,\ A\cap B=B\cap A. \tag{1.1.3}$$

(3) **分配律**

$$(A\cup B)\cap C=(A\cap C)\cup(B\cap C). \tag{1.1.4}$$

事实上, 若 $(A\cup B)\cap C$ 发生, 则 C 发生且 A 和 B 至少发生一个. 若 A 发生, 则 $A\cap C$ 发生. 若 B 发生, 则 $B\cap C$ 发生. 于是有

$$(A\cup B)\cap C\subset(A\cap C)\cup(B\cap C).$$

反过来, 若 $(A\cap C)\cup(B\cap C)$ 发生, 则 $A\cap C$ 和 $B\cap C$ 至少发生一个. 若 $A\cap C$ 发生, 则 A 和 C 都发生, 从而 $(A\cup B)\cap C$ 发生. 若 $(B\cap C)$ 发生, 则 B 和 C 都发生, 从而 $(A\cup B)\cap C$ 发生. 于是又有

$$(A\cup B)\cap C\supset(A\cap C)\cup(B\cap C).$$

这说明 (1.1.4) 成立.

另外, 读者还可以自己仿照以上说明, 得到

$$(A\cap B)\cup C=(A\cup C)\cap(B\cup C). \tag{1.1.5}$$

对于积 (交)、和 (并) 及对立运算还有

(4) **对偶律**

因为 $\overline{A\cup B}$ 表示 “不是 A 和 B 至少发生一个”, 亦即 A 和 B 都不发生, 所以有

$$\overline{A\cup B}=\overline{A}\cap\overline{B}. \tag{1.1.6}$$

同理, 可以说明

$$\overline{A\cap B}=\overline{A}\cup\overline{B}. \tag{1.1.7}$$

§1.2 概率的定义

前已指出, 概率论在研究随机现象的统计规律时, 是试图用一个数刻画一个随机事件发生的可能性的大小, 那么这个数存在吗? 对于一个事件, 这个数应如何规定呢?

1.2.1 概率的统计定义

首先, 大量的事实表明, 人们在对某一随机现象进行大量的重复观测时, 发现随着观测次数 n 的增加, 某一结果 A (事件) 发生次数 n_A 与 n 比值 (频率)

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

会在某一个介于 0 和 1 之间的数 p 附近摆动, 我们称 p 为事件 A 发生频率的稳定值. 这个 p 反映出事件 A 发生的可能性的大小, 称之为 A 发生的概率, 这便是**概率的统计定义**.

例如, 我们用同一种方式一次次地掷一枚均匀的骰子, 观察出现点数的情况. 若投掷次数 $n = 100$, 也许看不出各点出现的频率接近 1/6. 但若 $n = 1000$ 或 $n = 10000$, 就会发现各点出现的频率与 1/6 很接近. 在概率论发展的历史上, 曾有蒲丰 (Buffen) 和皮尔逊 (Pearson) 等人具体做过投均匀硬币的试验, 从中观测到随着投掷次数的增加, 出现正面和出现反面的频率越来越接近 50%.

概率的统计定义使人们相信可以用一个介于 0 和 1 之间的数来表示一个事件发生的可能性的大小, 即事件的概率是客观存在的, 但这种定义无法用来计算事件的概率, 因为试验次数多大才算合适, 无法确定. 另外, 所谓 "频率的稳定值", 只有在概率的公理化定义之后, 用后文介绍的 "大数定律" 才能明确阐述 (见 §5.1 推论 5.1.1). 于是, 人们着手探讨概率的其他定义, 比如针对某些特殊的随机现象, 给出计算概率的合理方法或公式, 这就是我们本节下面介绍的概率的古典定义和几何定义.

1.2.2 概率的古典定义

定义 1.2.1 设随机试验只有 $\omega_1, \omega_2, \cdots, \omega_n$ 等 n 个结果, 每次试验有且只有其中的一个发生, 每个结果发生的可能性大小相同. 则定义事件 A 发生的概率为

$$P(A) = \frac{n_A}{n}, \tag{1.2.1}$$

其中 n_A 为事件 A 中包含 $\{\omega_1, \omega_2, \cdots, \omega_n\}$ 中元素的个数.

我们称这种计算概率的数学模型为古典概型.

显然, 例 1.1.1 中 A, B, C 的概率正是按 (1.2.1) 来计算的.

例 1.2.1 袋中有 3 只白球 2 只红球, 现从袋中任取两只球, 试求以下各事件的概率.

(1) $A=\{$取得的两只球都是白球$\}$.

(2) $B=\{$取得的两只球都是红球$\}$.

(3) $C=\{$取得的球 1 只为白球 1 只为红球$\}$.

解 因为是 "任取" 两球, 所以取到 5 球中的任意两只球的可能性相同. 设想将 5 只球从 1 到 5 编号, 那么, 从中取出两只球共有 $\binom{5}{2}$ 种不同的结果, 所以 $n=\binom{5}{2}=\dfrac{5\times 4}{2!}=10.$

(1) 由于袋中有 3 只白球, 取出的两只白球必然在此 3 只球中抽取, 所以 $n_A=\binom{3}{2}=\dfrac{3\times 2}{2!}=3$, 于是

$$P(A)=\frac{3}{10}=0.3.$$

(2) 由于袋中有两只红球, 取出的两只红球就只能是这两只红球, 所以 $n_B=1$, 于是

$$P(B)=\frac{1}{10}=0.1.$$

(3) 由于袋中有 3 只白球两只红球, 取出的 1 只白球应在 3 只白球中抽取, 1 只红球应在两只红球中抽取. 所以, 1 只白球的取法有 3 种, 对白球的每种取法, 红球的取法有两种, 所以 $n_C=\binom{3}{1}\cdot\binom{2}{1}=3\times 2=6$, 于是

$$P(C)=\frac{6}{10}=0.6. \qquad \square$$

注: 这里及以后用专门的记号 $\binom{n}{m}$ 代替记号 $\mathrm{C}_n^m=\dfrac{n!}{m!(n-m)!}$.

例 1.2.2 试用古典概型解释划拳时叫哪个数最容易取胜.

甲乙两人划拳时, 每人等可能地伸出 0,1,2,3,4,5 个指头, 同时口中叫出 $0,1,2,\cdots,10$ 等 11 个数字, 每局若有一人叫出的数字等于两人伸出指头的和数就算该人赢, 两人都叫对或两人都叫错为平局.

我们用二维数组 (x,y) 来记甲, 乙伸出指头数, 共有 $6\times 6=36$ 种出拳方法, 而

{和数为 0}={(0,0)},

{和数为 1}={(1,0),(0,1)},
{和数为 2}={(2,0),(0,2),(1,1)},
{和数为 3}={(3,0),(0,3),(2,1),(1,2)},
{和数为 4}={(4,0),(0,4),(3,1),(1,3),(2,2)},
{和数为 5}={(5,0),(0,5),(4,1),(1,4),(3,2),(2,3)},
{和数为 6}={(5,1),(1,5),(4,2),(2,4),(3,3)},
{和数为 7}={(5,2),(2,5),(4,3),(3,4)},
{和数为 8}={(5,3),(3,5),(4,4)},
{和数为 9}={(5,4),(4,5)},
{和数为 10}={(5,5)}.
所以 "叫 5 而胜" 的概率最大,

$$P\{\text{叫 5 而胜}\} = \frac{6}{36} = \frac{1}{6}. \qquad \square$$

1.2.3 概率的几何定义

定义 1.2.2 设随机试验是往区域 Ω 里投点, 点落到 Ω 中某子区域 G 的可能性的大小只与 G 的度量大小有关, 而与 G 的形状和位置无关, 则定义

$$P\ (\text{点落到子区域 } G) = \frac{|G|}{|\Omega|}, \tag{1.2.2}$$

其中 $|\cdot|$ 表示几何度量, 它可以是长度, 面积, 体积等.

我们称这种计算概率的数学模型为几何概型.
显然, 例 1.1.2 中 A, B, C 的概率正是按 (1.2.2) 来计算的.

例 1.2.3 (约会问题) 两人相约在 7 点到 8 点间在某地会面, 先到者等候另一人 20 分钟, 过时即离去. 试求这两人能会面的概率.

解 依题意, 两人都在 7 点到 8 点间的任意时刻到达, 亦即在 [0,60] 的任意点到达, 设 x 和 y 分别为两人到达的时刻, 则两人到达的时刻为二维区域 $[0,60]\times[0,60]$ 内的所有点, 而两人能会面当且仅当

$$|x-y| \leqslant 20,$$

即能会面的到达时刻点 (x,y) 所形成区域的面积为, 边长为 60 的正方形面积, 减去 2 个直角边长为 40 的直角三角形面积, 所以

$$P\{\text{两人能会面}\} = \frac{60^2-40^2}{60^2} = \frac{5}{9}. \qquad \square$$

1.2.4 概率的公理化定义

前已指出, 概率的统计定义不能用来计算事件的概率, 更不便分析复杂随机现象的统计规律, 而概率的古典定义和几何定义又是基于某些具体的随机试验模型给出的, 并且该二类模型中都有“等可能性”这样一个苛刻的要求, 不能用来分析计算一般随机现象的统计规律, 这促使人们设法建立一套概率的公理化体系, 以便演绎地分析各类随机现象. 经过不断探索, 直到 20 世纪 30 年代, 以俄国数学家柯尔莫哥洛夫 (A. H. Колмогоров) 为代表的数学家建立了概率的公理化定义, 该定义的基本思想是把随机事件看作集合, 从而事件的和、积、对立及差等运算分别对应并、交、余和差等集合运算 (这一点, 读者已在例 1.1.3~例 1.1.5 中体会到了), 而把概率定义为集合的测度, 从而把概率论建立在测度论的基础上, 也使所有的讨论都在概率空间的框架下进行. 这里对应于前述随机试验, 随机事件和概率三个直观概念的分别是基本事件空间 Ω (或称样本空间), 事件域 $\mathcal{F}$ (或称事件 $\sigma-$ 代数) 和概率测度 P.

基本事件空间 (Ω)

定义 1.2.3 设 Ω 为一些事件构成的集合, 如果每次试验有且仅有 Ω 中的一个事件发生, 则称 Ω 为**基本事件空间** (或**样本空间**), 而称 Ω 中的事件为**基本事件** (或**样本点**)

从定义 1.2.3 中我们看到, 基本事件空间 Ω 是所有基本事件构成的集合 (全集), 基本事件之间互不相容, 且它们的和为必然事件. 其他事件则包含若干个基本事件, 也就是 Ω 的一个子集. 称一事件 A 发生, 当且仅当 A 中的一个基本事件发生. 所以, 称全集 Ω 为**必然事件**, 称空集 $\varnothing$ 为**不可能事件**.

例 1.2.4 (例 1.1.1 续) 记 $\omega_i=\{出现的点数为\ i\}, i=1,2,\cdots,6$, 则 $\Omega_1=\{\omega_1,\omega_2,\cdots,\omega_6\}$ 为一基本事件空间, 其中有 6 个基本事件. $A=\{\omega_1,\omega_3,\omega_5\}$, $B=\{\omega_2,\omega_4,\omega_6\}$, 且 $\Omega_2=\{A,B\}$ 也是基本事件空间, 其中只有两个基本事件. □

事件域 ($\mathcal{F}$)

事件域是人们研究随机现象时所感兴趣的事件集合, 由于事件经运算后仍为事件, 自然要求该集合对事件的运算封闭, 即其中的事件经事件运算之后仍然在该集合中, 这一性质可以概括为如下的定义 1.2.4.

定义 1.2.4 设 Ω 是基本事件空间, $\mathcal{F}$ 是 Ω 的一些子集所构成的集合 (类), 如果满足下列条件:

(1) $\Omega\in\mathcal{F}$.

(2) 若 $A\in\mathcal{F}$, 则 $\overline{A}\in\mathcal{F}$.

(3) 若 $A_n \in \mathcal{F}, n=1,2,\cdots$，则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$,

则称 $\mathcal{F}$ 为**事件域**, 并称 $\mathcal{F}$ 中的元素 (即 Ω 的某个子集) 为**事件**.

引进事件域定义的意义还在于对于同一个随机试验, 不同的观测者关心的事件可以不同.

例 1.2.5 (例 1.1.1 和例 1.2.4 续) 沿用例 1.1.1 和例 1.2.4 的记号, 则 $\mathcal{F}_1 = \{A \mid A \subset \Omega_1\}$, $\mathcal{F}_2 = \{\Omega_2, \varnothing, A, B\}$ 和 $\mathcal{F}_3 = \{\Omega_1, \varnothing, A, B\}$ 都为事件域. □

概率测度 (P)

概率测度是从我们日常生活中称重和量长度等度量方法抽象出来的. 比如, 两个西瓜的总重量等于各西瓜重量的和, 两个不重叠的线段的总长度等于各线段长度之和. 而称重的度量单位有千克, 磅等, 量长度的单位有米, 英尺等. 而概率测度则是事件域 $\mathcal{F}$ 中事件的一种度量, 一个事件在该度量下的大小, 代表该事件发生的可能性的大小.

定义 1.2.5 设 Ω 是随机试验的基本事件空间, $P(\cdot)$ 为定义在事件域 $\mathcal{F}$ 到实数集 $\mathbf{R}$ 的映射, 满足:

(1) 非负性 对任一事件 $A \in \mathcal{F}$, 有 $P(A) \geqslant 0$.

(2) 规范性 $P(\Omega) = 1$.

(3) 可列可加性 若事件 $A_1, A_2, \cdots$, 且两两互不相容, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

则称 P 为事件域 $\mathcal{F}$ 上的**概率测度**, 而称 $P(A)$ 为事件 A 的概率.

显然, 前面给出的概率的古典定义和几何定义, 都满足定义 1.2.5 中的 (1)~(3), 从而都是概率测度.

通常, 针对研究的某一随机现象, 都要先确定基本事件空间 Ω, 事件域 $\mathcal{F}$ 和概率测度 P, 将三者作为一个整体, 我们称 $(\Omega, \mathcal{F}, P)$ 为**概率空间**, 概率论的所有研究都是在这个空间上进行的.

概率的性质

由定义 1.2.5, 我们可以得到概率测度的一些常用的性质.

定理 1.2.1 设 $(\Omega, \mathcal{F}, P)$ 为概率空间, 则对 $\mathcal{F}$ 中的事件, 有

(1) $P(\varnothing) = 0$.

(2) 有限可加性　若 $A_1, A_2, \cdots, A_n$ 两两互不相容, 则

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

(3) $P(\overline{A}) = 1 - P(A)$.

(4) 单调性和可减性　若 $A \subset B$, 则有

$$P(A) \leqslant P(B), \quad \text{且 } P(B \backslash A) = P(B) - P(A).$$

(5) 加法公式

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

(6) 上 (下) 连续性　若 $A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \to \infty} P(A_n) \quad (\text{下连续性}).$$

若 $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$, 则

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \to \infty} P(A_n) \quad (\text{上连续性}).$$

证明　(1) 由于

$$\varnothing = \varnothing \cup \varnothing \cup \cdots \cup \varnothing \cdots,$$

从而, 由概率 P 的可列可加性, 有

$$P(\varnothing) = P(\varnothing) + P(\varnothing) + \cdots + P(\varnothing) \cdots,$$

这说明 $P(\varnothing) = 0$.

(2) 由于

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} A_i \cup \varnothing \cup \varnothing \cup \cdots \cup \varnothing \cdots,$$

从而, 由概率 P 的可列可加性, 有

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) + P(\varnothing) + \cdots + P(\varnothing) \cdots,$$

但 $P(\varnothing) = 0$, 所以 (2) 成立.

(3) 由于 $A \cup \overline{A} = \Omega$, 且 A 与 $\overline{A}$ 互不相容, 所以由 (2) 的结论有

$$1 = P(\Omega) = P(A) + P(\overline{A}).$$