数学概貌丛书

概率论概貌

王寿仁著

科学技术文献出版社

数学概貌丛书

概率论概貌

王寿仁　著

科学技术文献出版社

提 要

概率论是研究随机[illegible]门科学。概率论发端于十六世纪，当时还只用于研究赌博。但进[illegible]以后，概率论在生产实践中找到了非常广泛的应用，取得了飞速的发展。在概率论中，已大量运用了各种数学工具，形成了数学的一门独立分支。本书概要介绍概率论的基本内容，全书计分四章，即：概率定义、基本性质，极限理论，随机过程。本书基本上是介绍性的，通俗易懂，读者可从中对概率论这一数学分支有一个大致的了解。

本书的主要读者对象是大学数学系、物理系师生，旁学科的科学工作者，数学爱好者。

数学概貌丛书

概率论概貌

王寿仁 著

*

科学技术文献出版社出版

（北京市复兴路15号）

商务印书馆上海印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

开本 850×1156 1/32 印张 2 字数 47,000

1991年5月第1版 1991年5月第1次印刷

印数 1—2,500本

ISBN 7-5023-0662-5/O·52 定价：0.96元

目　录

一、概率定义

从本世纪三十年代起，概率论由一个很窄小的主题，发展成为一个宽广而与数学中许多其他领域有重要联系的分支．概率论是对客观存在的随机现象的数学抽象，正如几何学和力学一样，它同样也正确地反映了客观现实，具有广泛的应用．但是，概率论与几何学、力学在研究问题的方法上具有较大的差别，它所处理的不是决定性现象，而是随机现象，于是它有自己独特的概念和方法，在处理许多实际问题时具有独特的作用．

什么是“随机现象”呢？有这样一类现象，当人们对它加以观察或进行实验时，观察或实验的结果是许多可能结果中的某一个，而且这个或那个结果出现的机会是这个现象本身所固有的性质，我们把这种现象叫做**随机现象**，并把这些结果出现的机会的测度叫做**概率**．在初等水平上看这个问题，就是计算所出现的结果（叫做事件）的可能性大小．一个典型的例子是打桥牌，52 张牌分到四家，如果牌洗得彻底均匀，我们就可以计算某一家拿到 13 张不同点的牌的概率．计算这个概率，自然地会想到它应该等于下面这个比例：

$$\frac{\text{分 52 张牌时，某家拿到 13 张不同点的牌的可能情况的个数}}{\text{分 52 张牌时，这一家拿到 13 张牌的可能情况的个数}}.$$

接下来再求分子、分母的值，这只是计算个数的问题，属于组合数学的范畴了．由此可见，当我们谈概率时，总是把它结合着某种实

验来说的．实验(或观察)的结果称为**事件**．例如，在掷两个骰子的实验中，“总和为 6 点”、“双 6”、“两个都是奇数点”都是可能掷出的结果，所以它们都是事件．在这个实验里，最终可能的结果是：(1, 1)，(1, 2)，…，(1, 6)，(2, 1)，(2, 2)，…，(2, 6)，……，(6, 1)，(6, 2)，…，(6, 6)，一共三十六种．事件“双 6”，就是最后那个情形；而事件“总和为六点”，是指出现(1, 5)或(2, 4)或(3, 3)或(4, 2)或(5, 1)中的任何一个．由此可见，事件可以区分为**简单事件**与**复合事件**．例如“总和为 6 点”是个复合事件，它是由 5 个简单事件构成的；而“双 6”本身就是一个简单事件．

对以上例子加以抽象，我们看到：简单事件即是该实验的所有各种可能结果(这些简单事件叫做**样本点**，或简称为**点**)；复合事件可以分解为一些简单事件(也就是说，复合事件是由多个点组成的集合)．实验中的每一个不再可分解的结果就是点；由所有这些样本点构成一个总集合，称之为**样本空间**．由以上讨论可知，当谈到概率论时，首要的概念就是样本空间，其中的点为简单事件，这些点规定了该实验的所有可能结果，而且只有从实验的每个可能结果(点)都能够清楚地判定某一事件 E 是发生或不发生时，谈起 E 来才有确切的意义．判定“E 发生了”这个事件的那些点，构成一个集合，它完全确定事件“E 发生了”；反过来，含一个或多个点的给定集合 E 都可以当作一个事件，这个事件是否发生就由实验的结果(点)是否落在 E 中而定．在概率中，常说“事件 E 由某些点构成”，就意味着这些点便是在实验中能使 E 发生的那些结果．

1．样本空间

我们用 Ω 表示样本空间，其中的点用 ω 表示；事件是样本空间的子集合，用 A、B、…… 等表示．在一个实验中，谈到事件时，总是有不可能出现的事件，这反映在子集合上就是不包含 Ω 中任何点的子集，即**空集**，用符号 $\varnothing$ 表示．若 A 是一个事件，则“A 不

发生”也是一个事件，它由 Ω 中 A 以外的诸点组成，记作 A^c，称为 A 的补事件．实际问题中提出的各种各样的事件及其相互关系，均可反映为样本空间 Ω 的点组成的各种子集合及其相互关系上，因而，可以把概率论的语言与集合论的语言相互对照如下：

事件及其相互关系	集合及其相互关系	符 号
样本空间	空间	Ω
简单事件	点	ω
事件	集合	$A, B, \cdots$
不可能事件	空集	$\varnothing$
必然事件	整个空间	Ω
A 发生蕴含 B 发生	A 是 B 的子集	$A\subset B$，包含运算
A 与 B 同时发生	A 与 B 之交	$A\cap B$ 或 AB，乘运算
A 与 B 至少其中之一发生	A 与 B 之并	$A\cup B$，并运算
A 发生但 B 不发生	A 与 B 之差	$A\setminus B$，差运算
A 不发生（A 的补事件）	A 的补集	A^c
A 与 B 不同时发生（不相容）	A 与 B 不交	$A\cap B=\varnothing$

事件间的关系，可以表为图 1.

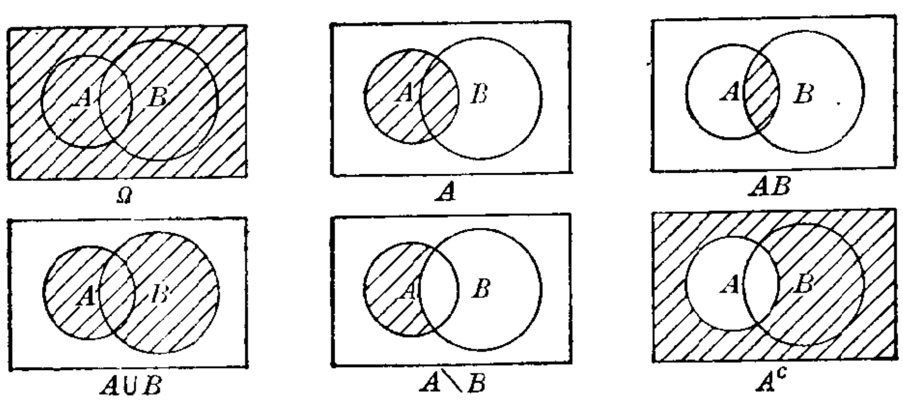

图 1

对于一串事件 A_1、A_2、A_3……，可定义两个新事件（即两个运算）：

(i) 属于全部给定集合的点所组成的集合，用 $\bigcap\limits_{i=1}^{\infty} A_i$ 表示，即

$$\bigcap_{i=1}^{\infty} A_i=\{\omega \mid \omega\in A_i,\ i=1,2,\cdots\cdots\}.$$

或记 $\bigcap\limits_{i=1}^{\infty} A_i=A_1A_2A_3\cdots\cdots$，称之为**交**．它表示 A_1、A_2、A_3、$\cdots\cdots$ 诸事件同时发生；

(ii) 由至少属于给定集中之一的诸点所组成的集合，用 $\bigcup\limits_{i=1}^{\infty} A_i$ 表示，即

$$\bigcup_{i=1}^{\infty} A_i=\{\omega \mid \omega\in A_i, \text{至少对某一个 } i\}.$$

或记 $\bigcup\limits_{i=1}^{\infty} A_i=A_1\cup A_2\cup\cdots$，称之为**并**．它表示事件 A_1、A_2、A_3、$\cdots$ 中至少有一个发生．我们还可以用很明显的极限形式表达以上两个事件

$$\bigcup_{i=1}^{\infty} A_i=\lim_{n\to\infty}\bigcup_{i=1}^{n} A_i, \qquad \bigcap_{i=1}^{\infty} A_i=\lim_{n\to\infty}\bigcap_{i=1}^{n} A_i.$$

显见，当 $n\to\infty$ 时，$\bigcup\limits_{i=1}^{n} A_i$ 不降 $\left(\text{即 } \bigcup\limits_{i=1}^{n} A_i\subseteq\bigcup\limits_{i=1}^{n+1} A_i\right)$，而 $\bigcap\limits_{i=1}^{n} A_i$ 不增．所以，前者涨到 $\bigcup\limits_{i=1}^{\infty} A_i$，后者缩到 $\bigcap\limits_{i=1}^{\infty} A_i$．

此外，对事件 (A_i) 为有穷个或无穷个事件时，都成立以下的运算：

$$\left(\bigcup_i A_i\right)B=\bigcup_i (A_iB),$$

$$\left(\bigcap_i A_i\right)^c=\bigcup_i A_i^c.$$

2. 古典概率

古典概率讨论的对象，限于实验所有可能结果为有穷个等可能的情形．这时，样本空间 Ω 由 n 个简单事件组成．若事件 E 由

m 个简单事件组成, 由直觉可以计算 E 的概率 $P(E)=\frac{m}{n}$, 这就是法国数学家拉普拉斯(P. S. M.de Laplace, 1749~1827)对古典概率作的定义. 让我们来计算打桥牌时, 分到西家一副 13 张不同点的牌之概率. 首先, 分到西家 13 张牌所有可能情况的副数等于 $C_{52}^{13}=\frac{52\cdot 51\cdot\cdots\cdot 40}{13!}=635013559600$ 副; 分到西家 13 张不同点的牌所有可能情况的个数即从 4 张 A 中抽一张, 4 张 K 中抽一张, ……, 所以总共可抽得 4^{13} 副不同点的牌. 从而, 所欲计算的概率等于

$$P=\frac{4^{13}}{C_{52}^{13}}\approx 0.0001057.$$

下面再举一个历史上有名的"得分问题". 甲、乙两人各出同样的赌注, 用扔硬币为赌博手段. 若正面朝上(记为"+"), 甲得 1 分, 乙不得分; 若反面朝上(记为"−"), 乙得 1 分, 甲不得分. 谁先得到事前约定的分数, 谁就赢得全部赌注. 当进行到甲还差两分, 乙还差三分才能达到约定分数时, 他们不愿继续赌下去, 问这时如何公平分配全部赌注? 解此问题, 显然, 为确保能分出胜负, 最多需要再扔 4 次, 这 4 次的所有可能结果为: {(++++), (+++−), (++−+), (++−−), (+−++), (−+++), (+−+−), (−++−), (−−−+), (−−−−), (+−−+), (−+−+), (−−++), (+−−−), (−+−−), (−−+−)}, 其中使甲获胜(即至少两个"+"的情形)有 11 种, 使乙获胜(即至少三个"−"的情形)有 5 种, 故甲胜的概率为 $\frac{11}{16}$, 乙胜的概率为 $\frac{5}{16}$, 因而甲应得全部赌注的 $\frac{11}{16}$, 乙应得 $\frac{5}{16}$. 实际上, 前四种情形表示只需扔两次就可分胜负, 往下的六种情形表示需扔三次即可分胜负, 最后的六种表示非要扔四次才能决定胜负. 但如果这样考虑, 就不利于用古典概率进行计算了.

计算古典概率, 可以像上例那样, 采用穷举法数清事件里面所

含简单事件的个数．借助于组合计算，可以简化计算过程．如上例中，"甲胜"事件所含简单事件个数为 $C_4^4+C_4^3+C_4^2=11$．需要强调的是，这里应十分注意"等可能"这个性质，即每个简单事件具有相同的概率．历史上，法国数学家达朗贝尔(J. L. R. D'Alembert, 1717～1783)在考虑掷两个硬币的实验中，曾错误地认为"出现一个正面朝上一个反面朝上"的概率为 $\frac{1}{3}$．事实上，此实验有 $\{(++), (+-), (-+), (--)\}$ 四个等可能的简单事件，而事件一反一正含 $(+-)$、$(-+)$ 两个简单事件，其概率应为 $\frac{1}{2}$．

在古典概率模型中，可把事件 A 的概率 $P(A)$ 看成事件 A 的函数，它具有下列三个性质：

(i) 对于任意事件 A，有 $0\leqslant P(A)\leqslant 1$；

(ii) 必然事件 Ω 的概率 $P(\Omega)=1$；

(iii) 若 A、B 不相容，其中至少一个发生即 $A\cup B$（或记为 $A+B$，对不相容的情形）的概率为 $P(A+B)=P(A)+P(B)$．

性质(iii)是计算概率的重要法则．例如，有一批产品，总数为 1000 件，其中 10 件为废品，从这 1000 件中随意抽出 20 件，求这 20 件中有废品的概率．用 A 表示这一事件，如直接计算 A 的概率，就要依次计算 20 件中恰有 1 个、2 个、……直至 10 个为废品的概率，再求这些概率之和．但 A 的补事件 A^c 是这 20 件中无一废品．由性质(ii)、(iii)知 $P(A)+P(A^c)=P(\Omega)=1$，于是 $P(A)=1-P(A^c)$，而计算 $P(A^c)$ 是很容易的，它等于 $\frac{C_{990}^{20}}{C_{1000}^{20}}$，故 $P(A)=1-\frac{C_{990}^{20}}{C_{1000}^{20}}$．

3. 几何概率

在建立古典概率的同时，人们就注意到只考虑有穷个简单事件构成的样本空间是不够用的．把等可能思想应用到含无穷个点

的样本空间，就产生了几何概率．其基本思想可以表述为：设 Ω 是平面上一个可求面积的区域，其面积记为 $\mu(\Omega)$．设 Ω 中的子集 A（即事件）的面积为 $\mu(A)$．随意扔一个点到 Ω 中，按照等可能性的思想，可以认为落入 Ω 中子集 A 的概率等于这个子集的面积与 $\mu(\Omega)$ 之比，即

$$P(A)=\frac{\mu(A)}{\mu(\Omega)}.$$

下面的蒲丰(G. L. de Buffon, 1707～1788)投针问题是一个应用几何概率的典型例子．设平面上有一族平行线，每相邻两条之间的距离为 2 单位．取单位长的针一枚，随意把它扔到平面上，求针与直线相交的概率．用 x 表示针的中点到离它最近的一条平行线的距离，用 θ 表示针与此平行线的夹角(如图 2)．(x, θ) 完全决定针所落的位置．针的所有可能位置为 $0\leqslant x\leqslant 1, 0\leqslant\theta<\pi$（即图 3 所示的矩形里的每一个点），矩形的面积为 π．针与直线相交的充分必要条件是 $x\leqslant\frac{1}{2}\sin\theta$，这一不等式相应于图 3 中的阴影部分，其面积等于 $\frac{1}{2}\int_0^\pi \sin\theta\, d\theta=1$．用 A 表示事件“针与直线相交”，则 $P(A)=\frac{1}{\pi}$．利用这一公式，重复 n 次投针，计数它与直线相交的次数 m，则当 n 很大时，相交的频率 $\frac{m}{n}$ 可以作为 $P(A)=\frac{1}{\pi}$ 的近似(参看下节的讨论)．虽然这样做既费时，又不精确，但这一思想有可取之处，特别在有了电子计算机之后，可以用电子计算机模拟投针试验，计算 π 的近似值，这就显出其优越性了．许多与此

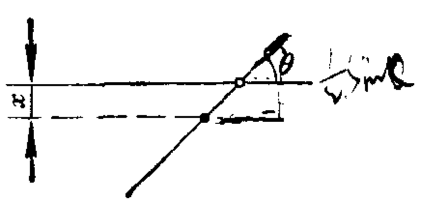

图 2

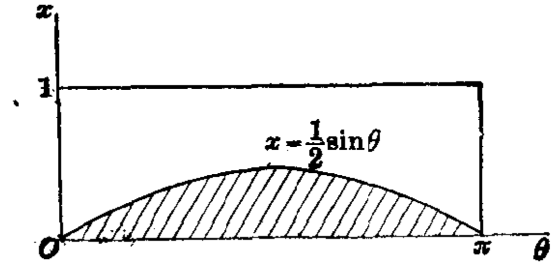

图 3

例相象的问题，表面上看来和概率毫不相干，但可用几何概率加以解决.

4. 概率的一般定义

数学科学的发展，总是结合着实际，把具体的对象扩展到抽象的对象，把特殊的事物推广到一般的事物. 首先我们来考虑等可能性的假设. 因为利用等可能性曾经遇到了悖论，从而引起人们注意到在一般情形下，等可能性这个前提是否必要. 下面先介绍这个悖论(贝特朗(Bertrand)悖论)：

在半径为1的圆内随意取一弦，计算此弦的长超过圆内接等边三角形边长 $\sqrt{3}$ 的概率.

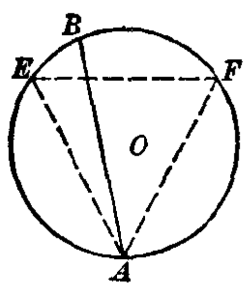

图 4

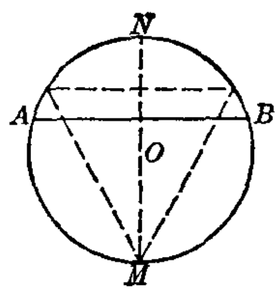

图 5

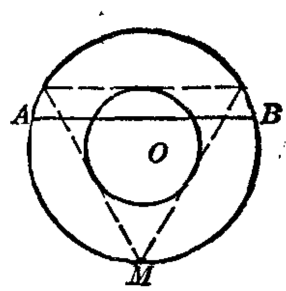

图 6

解此题有三种考虑方法：

如图4，AB 为弦. 不失一般性，把弦的一端 A 固定，令点 B 在圆周上移动. 显然，当 B 点在 $\overparen{EF}$ 上时，AB 的长超过 $\sqrt{3}$，而 $\overparen{EF}$ 为圆周的 $\frac{1}{3}$，故所求概率为 $\frac{1}{3}$.

如图5，弦 AB 的长与它与圆心 O 的距离有关，而与它的方向无关. 因此，可设它垂直于某个直径，不妨设此直径为 MN. 由图5看出，当且仅当弦与圆心的距离小于 $\frac{1}{2}$ 时，弦的长大于 $\sqrt{3}$. 因此，所求概率为 $\frac{1}{2}$.

如图6，弦的长被它的中点所决定，从图中可以看出，当且仅

当中点在半径为 $\frac{1}{2}$ 的同心圆内时，弦长大于 $\sqrt{3}$. 但这个同心圆的面积为大圆面积的 $\frac{1}{4}$, 故所求概率为 $\frac{1}{4}$.

这个悖论的根源, 在于取弦时采用了不同的等可能性, 图 4 是把 B 点在圆周上看成有等可能性, 图 5 是把弦中点在直径上看成有等可能性, 而图 6 则把弦的中点在圆内看成有等可能性. 由此例看出, 等可能性引起了怪现象!

在一些具体的例子如掷骰子实验中，人们还对事物加以理想化, 如假设骰子为一个绝对对称的正六面体. 但是, 现实中这种纯而又纯的事物是不存在的. 更深一步, 即使存在绝对对称的骰子, 但由于各人的掷法不同, 也会破坏等可能性. 此外, 我们对每个点出现的概率规定为 $\frac{1}{6}$, 这是出于人们的常识. 但对常识需要有直观的解释, $\frac{1}{6}$ 应该给出每个点大概要出现的机会的一个度量，而这个度量在经验上与多次掷骰子时某个点出现的频率相联系. 更一般地，令 E 表示“所出现的结果小于 5 点”, 设我们重复掷 n 次骰子, 令 $N_n(E)$ 表示这 n 次结果中 E 出现的次数, 于是 E 出现的频率 $Q_n(E)=N_n(E)/n$, 我们有理由取 $Q_n(E)$ 作为事件 E 出现的度量. 当然 $Q_n(E)$ 与 n 有关, 而且当 n 增多时, $Q_n(E)$ 不断地在摆动, 甚至摆动很大. 但若 n 趋于无穷, 随着 n 的增大, 序列 $Q_n(E)$ 是否能在一个固定值的很小范围内摆动呢? 也就是 $Q_n(E)$ 是否稳定到一个固定值呢? 深究起来，我们不能答覆此问题, 因为这是一个永不完结的实验. 但在数学上, 我们可以理想化而规定这个极限存在，并记其极限为 $Q(E)$，这就是 $Q(E)=\lim\limits_{n\to\infty}Q_n(E)$，并称 $Q(E)$ 为事件 E 的直觉极限频率. 但这个极限, 更深究起来, 仍然与所取的具体的实验序列有关, 即对第一个人作的实验(假定可以作)与第二个人，第三个人……所作的实验来说, 不能保证他们能得到相同极限. 然而, 从实际的直觉, 我们需要有一个衡量事件 E 出现的机会的度量，而这个度量将比只作一次 $n\to\infty$ 的实验所得

的记录要含有更多的东西．一个能够维持下去的以频率为基础的理论需要假定：对一切"相类似"的序列而言，上述极限都是一样的等于 $Q(E)$．但这一假设里，"相类似"的含义还是不明确的．所以，这种办法只在一定范围内可以有效，而且，在直觉的常识范围内它具有诱惑力．这里，我们不去穷追此问题．下面将对事件 E 的概率给以合理的定义，在此定义下所得到的概率理论中可以证明一个基本定理：在某些条件下，事件 E 的频率 $\frac{N_n(E)}{n}$，对实际所有能考虑到的序列而言，它的极限是存在的，且等于事件 E 的概率．这个定理就是**大数定理**，它是一切实验科学应用概率论的基础，而且，在理论上验证了以频率来解释概率的直觉观念．

设 Ω 为样本空间，其中的事件具有样本空间里关于事件间的一切关系及运算．我们定义每个事件 E 的概率 $P(E)$ 为 0 与 1 之间的实数，它度量 E 发生的机会．这种度量，犹如力学中的质量的度量或几何学中距离的度量，是事件所固有的性质．至于这些度量等于什么，怎样测出来的，用不着再作其它假定．但当人们真需要知道概率数值时，就需要人们的经验及统计方法来定出．为了演算和推理，概率或称概率测度应具有初等概率所具有的三个性质，我们把这三个性质作为规定（公理）：

（i）对每个事件 $E(\subset\Omega)$，$P(E)\geqslant 0$；

（ii）对任意两个不相容事件 E_1、E_2，它们的并（即至少有一个发生的事件）E_1+E_2 的概率满足：

$$P(E_1+E_2)=P(E_1)+P(E_2);$$

(iii) $P(\Omega)=1$.

由以上三条公理，可以推出以下的性质：

(iv) 对任意事件 E，$P(E)\leqslant 1$.

这是因为，E 与其补事件 E^c 之并等于样本空间 Ω，即 $E+E^c=\Omega$；于是，由 (ii) 知 $P(E)+P(E^c)=P(\Omega)=1$，从而，由 (i) 知 $P(E)=1-P(E^c)\leqslant 1$.

(v) 对任意两个事件 E_1、E_2, 且 $E_1 \subset E_2$, 则有

$$P(E_1) \leqslant P(E_2), \quad P(E_2 \setminus E_1) = P(E_2) - P(E_1).$$

这是因为$E_2 = E_1 + (E_2 \setminus E_1)$, 于是 $P(E_2) = P(E_1) + P(E_2 \setminus E_1) \geqslant P(E_1)$.

(vi) 对任意有穷个事件 E_1、E_2、…、E_n 且两两不相容, 有下列的有穷可加性:

$$P(E_1 + E_2 + \cdots + E_n) = P(E_1) + P(E_2) + \cdots + P(E_n).$$

(vii) 对任意有穷个事件 E_1、E_2、…、E_n, 恒成立下列布尔不等式:

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) \leqslant P(E_1) + \cdots + P(E_n).$$

这是因为, 当 $n=2$ 时, $E_1 \cup E_2 = E_1 \cup E_1^c E_2$, 对右边可用公理(ii), 得 $P(E_1 \cup E_2) = P(E_1) + P(E_1^c E_2)$; 但 $E_1^c E_2 \subset E_2$, 再用公理(v), 即得 $P(E_1 \cup E_2) \leqslant P(E_1) + P(E_2)$. 对一般的 n, 可用数学归纳法来证.

(viii) 对任意两个事件 E_1、E_2, 有

$$P(E_1 \cup E_2) + P(E_1 \cap E_2) = P(E_1) + P(E_2).$$

这是因为, $P(E_1 \cup E_2) = P(E_1) + P(E_1^c E_2)$, 但 $E_1^c E_2 = E_2 \setminus E_1 \cap E_2$, 故有 $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

为了应付更广泛的情形, 我们将把公理(ii)加以强化. 这种强化是符合客观实际的, 而且由它推得的结论也是合理的. 公理(ii)及其直接推论(vi)可强化为“**可列可加性**”公理:

(ii*) 对两两不相容的事件的无穷序列 E_i(其中 $i=1,2,\cdots$), 有

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

注意, $\bigcup_{i=1}^{\infty} E_i$ 仍是事件, 它是诸 E_i (其中 $i=1, 2, \cdots$)中至少一个发生的事件.

如果 Ω 为有穷个点构成的, 则(ii*)化为(ii). 但是, 重要的是可列可加性不能由有穷可加性(vi)令 $n \to \infty$ 而得到. 现在让我们

看一看：有穷可加性可写为

$$P\left(\bigcup_{i=1}^{n} E_i\right)=\sum_{i=1}^{n} P(E_i), \quad E_i \text{ 两两不相容} (i=1, \cdots, n).$$

当 n 增加时，右方是非降的，而且其和不超过 1，故右方有极限 $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n} P(E_i)=\sum\limits_{i=1}^{\infty} P(E_i)$. 于是，由上式得

$$\lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right)=\sum_{i=1}^{\infty} P(E_i).$$

欲使此等式化为可列可加性，需成立下列关系：

$$\lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right)=P\left(\bigcup_{i=1}^{\infty} E_i\right)=P\left(\lim_{n\to\infty}\bigcup_{i=1}^{n} E_i\right).$$

但是，此式两端需要运算“$\lim\limits_{n\to\infty}$”与“P”可以互换. 如果读者学过微积分，就会知道这种互换不是随便可以进行的，而是有条件的.

二、基 本 性 质

1. 事件的独立性

独立性是概率论特有的概念. 考虑“甲扔硬币，乙掷骰子”的实验，求事件“硬币出现正面、并且骰子出现 5 或 6 点”的概率，若把“硬币出现正面”记作事件 A，“骰子出现 5 或 6 点”记作事件 B，则欲求的概率就是事件 $A\cap B$ 的概率. 整个实验的简单事件的总个数为 12，而且是等可能的. A 所含简单事件的个数为 6，B 所含简单事件的个数为 4（即，(正，5)，(反，5)，(正，6)，(反，6)），而 $A\cap B$ 所含简单事件的个数为 2，于是 $P(A)=\frac{6}{12}=\frac{1}{2}$，$P(B)=\frac{4}{12}=\frac{1}{3}$，$P(A\cap B)=\frac{2}{12}=\frac{1}{6}$. 由此可知 $P(A\cap B)=P(A)P(B)$. 对于任意两个事件 A 和 B，如果它们各自发生的概

率与它们同时发生的概率满足 $P(A\cap B)=P(A)P(B)$，则称它们为相互独立的．结合上例，独立性这一定义与直观是一致的，因为甲和乙作实验是单独进行而互不影响的．

对于 n 个事件 A_1、A_2、…、A_n，$n\geqslant 2$，如果对所有可能的 $1\leqslant i_1<i_2<\cdots<i_l\leqslant n$（其中 $2\leqslant l\leqslant n$），都成立

$$P\left(\bigcap_{j=1}^{l} A_{i_j}\right)=\prod_{j=1}^{l} P(A_{i_j}),$$

则称这 n 个事件相互独立．以上的关系式需要对 $l=2, \cdots, n$ 和任意 $1\leqslant i_1<i_2<\cdots<i_l\leqslant n$ 都成立，缺一不可．这可由下面的例子说明：A、B、C 三个事件是两两相互独立，但非相互独立．考虑一个正四面体，一面为红色，一面为蓝色，一面为黄色，剩下的那面为红、蓝、黄三色．以 A、B、C 分别表示掷四面体时，朝下的一面出现红、蓝、黄的事件．于是 $P(A)=P(B)=P(C)=\frac{1}{2}$，且 $P(AB)=P(AC)=P(BC)=\frac{1}{4}$．由此知 A、B、C 两两相互独立．但 $P(ABC)=\frac{1}{4}$，$P(A)P(B)P(C)=\frac{1}{8}$，不满足 $n=l=3$ 所要求的条件，所以事件 A、B、C 不相互独立．

［例］（伯努利(Jacob Bernoulli, 1654～1705)实验）对相同条件下重复的同一实验，独立性是很有用的．例如重复扔硬币 n 次，$n\geqslant 2$，则整个重复试验的结果是由“正面”、“反面”组成的长为 n 的序列，我们常常把“正”、“反”分别量化为“1”，“0”，或“1”、“−1”．现采取前面一种量化，于是重复扔硬币 n 次的实验结果是由“0”、“1”组成的长为 n 的序列．例如 $n=9$，(110010100) 就是一个结果．对重复扔 n 次，一共有 2^n 个这样的序列．由对称性假设及独立性，每一个简单事件即出现某一“0”、“1” n 序列的概率为 $\frac{1}{2^n}$．如果把对称性假设去掉（有这类的实际问题），并设出现 1 的概率为 $p\left(\neq\frac{1}{2}\right)$，而出现 0 的概率自然是 $1-p$（记为 q），现在用数学公式来表示，设 X_i 为这 n 次重复试验中的第 i 次的结果，