概率论入门

杨宗磐 著

科学出版社

概 率 论 入 门

杨宗磐 著

科 学 出 版 社

1 9 8 1

内 容 简 介

本书是作者在讲授概率论的讲稿基础上编写而成的．内容主要有：概率的古典和几何定义；Колмогоров 公理系；随机变数的数字表征；大数定律和极限定理；随机过程的一般理论；马氏过程和平稳过程等．

本书结合许多物理模型说明概念，深入浅出．作者还指出了古典概念和现代概念的联系以及所使用的数学工具的多样化．每节后附有习题，以使读者进一步加深理解内容．

读者对象是大专院校数学系高年级学生、教师和数学工作者．

概 率 论 入 门

杨宗磐 著

责任编辑 刘嘉善

科学出版社出版

北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1981 年 8 月第 一 版 开本：850×1168 1/32

1981 年 8 月第一次印刷 印张：12

印数：0001—20,200 字数：312,000

统一书号：13031・1588

本社书号：2184・13—1

定价： 2.20 元

整理者的话

本书是杨宗磐先生早在六十年代初就已脱稿准备出版的．但是由于种种原因一直未能如愿．时光流逝，不觉已过十五年，杨先生也已于 1976 年 3 月去世．

为了更好地适应目前我国的情况，在不影响本书宗旨的前提下，我们对原稿重新作了整理，在内容和措词上作了一些改动．希望这样能使读者在阅读时更便于领会作者的原意．

吴　荣

序

为了表达对已故杨宗磐同志的怀念，也为了尊重他的遗愿——叫我在他写的这部书中做一点工作，我写一篇稍长一些的序，希望有助于清除有关概率论的一些糊涂概念.

当用雷达跟踪像飞机一类目标时，我们量测到的是目标的斜距、高低角和方位角. 对于这三个量的每次量测都不可避免地带有误差. 影响量测精度的因素非常多：有来源于安装着雷达的载体的(例如飞机、船)；有来自目标的(例如目标闪烁噪声等)；有来自电磁波传播的(传播误差、多路误差等)；有来自周围环境的(大气折射、大气中电磁现象等)，等等. 这些因素本身又都不断地变化着，而影响它们变化的又有许多因素. 确定雷达每次量测中误差的大小是非常复杂的，因而不能满足实时需要. 雷达量测是这样，我们使用的任何一种量测仪表也都是这样.

又当考虑飞机飞行状况时，不能不考虑阵风的影响. 确定各时刻阵风的大小与方向的因素也非常多，从而要确定阵风在每个时刻的大小与方向而不带任何误差，也极难做到，更不用说满足实时需要了. 关于阵风是这样，关于大气密度的变化，海洋水流的变化等等，也都是这样.

一些教科书中常谈到影响以上这类大量现象的因素无法完全知道. 问题不仅在于能不能知道雷达每次量测误差的大小或每时刻阵风的方向、大小，问题还在于对工程技术人员来说，需要不需要知道这些，或至少在不知道这些时有无办法有效地使用雷达量测数据或设计很好的飞行控制系统.

值得注意，每次雷达量测并不是孤立事件；每时刻阵风的方向与大小也不是孤立的. 在使用雷达跟踪目标时，我们取得成百次量测数据. 如果研究一下这种多次观察中看到的现象，在表面杂乱无章中却存在着一定的规律性. 科学的任务正是透过一切表面

的偶然性揭示过程的内在规律性.

有一种极端看法,根本否认偶然性.如果说要确定出雷达每次量测中误差的大小非常复杂,那末,能否期待有朝一日把雷达量测的每次的噪声来源及其大小弄得一清二楚呢?恩格斯早在谈到决定论观点时就曾批评过那种企图对一个豌豆荚中究竟有几粒豌豆也用因果关系加以说明的人,认为这种"也还是没有从神学的自然观中走出来"(《自然辩证法》,人民出版社1971年版196~197页).显然致力于探索雷达个别一次量测几十种噪声源在这特定时刻特定环境下的具体情况,从这些噪声的因果连锁方面探究这次量测噪声的大小,这不是什么科学,并不比研究一个豌豆荚中有几粒豆的人更伟大!我们关心的是在每次观察或实验中带有偶然性而在大量重复性观察中呈现出一种特定的规律性的现象,即我们称做随机现象.雷达量测噪声、阵风都属于这种随机现象.概率论就是从量的关系方面研究随机现象的数学学科.

既然随机现象是在大量重复性实验或观测的基础上谈的,我们把这种实验或观测抽象出来,成为实验$\mathscr{E}$的概念.在每次实验结果(或实验$\mathscr{E}$的每次实现)a中,一特定事件A或是发生,或是不发生,这里并没有任何不确切性.例如$\mathscr{E}$表示用雷达跟踪目标测得目标位置数据的整个过程.实验$\mathscr{E}$的每次实现a就表示雷达的一次量测.某特定的事件,比如说,指这一次量测误差不超过某定量.我们常把事件A等同于实验$\mathscr{E}$中使事件A发生的那些实现的集合,于是事件的演算等同于必然事件Ω的一些子集(即事件)的演算.特别如果$E[x<\alpha]$表示所观测的量x的值小于给定的实数α这一事件——比如说,雷达测距误差小于2米这一事件,而$E[x\leqslant\alpha]$表示所观测的量x的值不超过α这一事件.为了在事件演算中包括诸如下列的等式:

$$E[x<\alpha]=E[x\leqslant\alpha-1]\vee E\left[x\leqslant\alpha-\frac{1}{2}\right]\vee\cdots$$
$$\vee E\left[x\leqslant\alpha-\frac{1}{n}\right]\vee E\left[x\leqslant\alpha-\frac{1}{n+1}\right]\vee\cdots,\qquad(1)$$

在集演算中应包括集的可数并与可数交的运算：$\bigcup\limits_{n=1}^{\infty} A_n$与 $\bigcap\limits_{n=1}^{\infty} A_n$. 这就是说，所考虑的事件是必然事件 Ω 的一些子集所构成的 σ-代数(或简称系 (tribu)). 于是讨论随机现象时，我们从两个基本概念出发：Ω 与 Ω 的一些子集的系 $\mathscr{F}$，习惯上合记成 $(\Omega, \mathscr{F})$，称做“可测空间”.

此外，要对一随机事件发生可能性大小有一数量的量度. 如果考察一事件 A，我们反复作多次实验 $\mathscr{E}$，取得 $\mathscr{E}$ 的一串实现 $a_1, a_2, \cdots, a_n$，那末在这 n 次实现中 A 事件发生的次数能表达成

$$\sum_{\substack{a_k \in A \\ 1 \leqslant k \leqslant n}} 1. \tag{2}$$

这个和号指每有一个 $a_k \in A$，和中就有一个 1. 事件 A 在实现串 $a_1, a_2, \cdots, a_n$ 中发生的频率，指在这串中 A 发生的次数(2)与实验实现总次数 n 之比，即

$$\frac{1}{n}\Big(\sum_{a_k \in A} 1\Big) = \frac{1}{n} \sum_{k=1}^{n} \chi_A(a_k), \tag{3}$$

这里 χ_A 表示集 A 的示性函数：

$$\chi_A(a) = \begin{cases} 1, & \text{如果 } a \in A, \\ 0, & \text{如果 } a \notin A. \end{cases}$$

从实践中人们获得这样的认识：比值 (3) 随 n 变大而愈来愈环绕着某个确定数值而起伏. 在一定场合所能实现的实验次数总是有穷的，但如多次重复作一串实验 $\mathscr{E}$，对于每串象上述那样的实验实现，比值 (3) 总环绕某确定值 p 而起伏，用归纳法就认为事件 A 发生的可能性有一客观的，即独立于作实验 $\mathscr{E}$ 的人的主观意志的量度 p.

因此，当研究随机现象时，不但要给定可测度间 $(\Omega, \mathscr{F})$，还要考虑事件 A 发生可能性的量度. 这个量度就是定义在 $\mathscr{F}$ 上的函数 $p(\cdot)$；对于每个 $A \in \mathscr{F}$，$p(A)$ 是一确定的实数，叫做事件 A 的概率，而函数 $A \to p(A)$ 叫做 $(\Omega, \mathscr{F})$ 上的概率测度. 带有概率测度 $p(\cdot)$ 的可测空间 $(\Omega, \mathscr{F})$ 叫做概率空间，记成 $(\Omega,$

$\mathscr{F}, P)$. 从(2)(3)可以看出，为了反映事件A发生可能性大小的量度，概率测度$P(\cdot)$应有如下的属性:

1) $P(\Omega)=1$;

2) $P(A)\geqslant 0\ \ \forall A\in\mathscr{F}$;

3) $A_i\in\mathscr{F}, i=1,2,\cdots,n,\cdots$, 且 $A_i\wedge A_j=\varnothing\ (i\neq j)$ [即诸 A_i 是两两互斥事件] $\Longrightarrow$

$$P\left(\bigvee_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i). \tag{4}$$

这里3)对于有穷多相加项从(2)(3)直接看出,即有穷加法性从实践能推断出来,但可数无穷加法性是为了考虑到事件的属性(1)而从有有穷加法性引伸出来. 1) 2) 3) 就是关于概率测度的公理系统,也可以说,是概率测度$p(\cdot)$的隐定义,因为1) 2) 3)并没有正面说什么是概率测度，而是把实践经验中关于概率的认识提炼成这三条主要的数量关系.

在实际问题中考虑的并不直接是集A，而往往是随机现象的一些数量特征. 为了定量地刻划随机现象，每次实验实现确定出的是一些数量的值. 例如一次雷达观测获得的是三个数值: 目标的斜距、高低角、方位角. 一般说,实验的实现确定出n个数值,看作n维实空间$\mathbf{R}^n$中的一点(或n维矢量).每次实验的实现 $\omega\in\Omega$ 确定出来的量$x(\omega)$是$\mathbf{R}^n$中的元,即$x(\cdot):\Omega\to\mathbf{R}^n$. 在研究随机现象时所关心的是如下一类事件:

$$x(\omega)\leqslant\alpha;\quad x(\omega)=\alpha;\quad x(\omega)<\alpha$$

等等. 例如雷达测距所关心的是一次量测ω之下测距误差$x(\omega)$不超过α米这样的事件: $x(\omega)\leqslant\alpha$. 因此，在取定了概率空间$(\Omega,\mathscr{F},P)$之后，所考察的量$x(\omega)$不是任意函数$x(\cdot):\Omega\to\mathbf{R}^n$，而应满足条件 $\{\omega\in\Omega\,|\,x(\omega)\leqslant\alpha\}\in\mathscr{F},\forall\alpha\in\mathbf{R}$. 满足这种要求的量$x$叫做随机变量;从函数$\omega\longmapsto x(\omega)$来说，随机变量$x$正是$(\Omega,\mathscr{F})$上的可测函数.

上面我们主要说明了概率论的基本概念以及其公理系统怎样在人类实践的基础上,在感性认识的基础上,经过人的思维活动的

加工而提炼出来．若干年来颇有一些人怕谈公理系统，似乎一讲公理系统就必然与纯演绎体系联系起来，怕被人批判为先验论等等．这里强调了概率论的公理系统由在实践中积累的感性认识提炼出来．一旦归纳成几条公理的工作完成，概率论就能从它的公理系统推演出来．但这不过是写书的一种叙述方法，与先验论的哲学思想并无关系．为此，我们最好看一看马克思的说明（见《资本论》第一卷第二版跋)：

“当然，在形式上，叙述方法必须与研究方法不同．研究必须充分地占有材料，分析它的各种发展形式，探寻这些形式的内在联系．只有这项工作完成以后，现实的运动才能适当地叙述出来．这点一旦做到，材料的生命一旦观念地反映出来，呈现在我们面前的就好象是一个先验的结构了．”

作为一部数学的概率论的书，利用公理系统作严格推导是一种好的叙述方法；而当我们把概率论的基本概念、它的公理系统的提炼过程加以细致的分析时，外表的先验结构所隐藏的实践来源也就清楚了．杨宗磐同志的这部书中大量的物理学的例当然会更好地帮助读者了解概率论中一些基本概念的实践来源．

这部书与一般教科书、入门书有很大的不同．著者在他对概率论的数学基础的深刻理解之下举了大量正面与反面例子，并以细致的分析帮助读者对于这门学科获得确切的知识．这是本书突出的特色．相信学过概率论基础课程的人通过阅读这部书会加深对概率论的理解．

关肇直

1980年3月10日于北京

作 者 序

作者五十年代在南开大学讲授概率论课．本书是在当时的讲稿的基础上改写而成的．

作者写这本书的主观愿望是为已经阅读过概率论基础教材，具备数学系三年级专业知识的读者，特别是自学的读者，在进一步学习时提供一些参考材料．作者根据自己的学习经验，在本书中力图防止把概率论当成五个主要孤立部分：古典及几何定义，Колмогоров 公理系，增函数列，几个偏微分方程，又像又不像实变函数的一部分．概率论处理的是随时间变动的随机现象，即过程（包括独立随机变数列）；集论的诞生使处理变量的微积分学有了坚实的基础，概率论里 Колмогоров 公理系的提出，相当于集论的诞生，有关概率的问题只有纳入 Колмогоров 公理系才算有了正当的归宿．本书的内容不多，主要是：1. 指出初步了解决定论规律及概率性规律的参考书；2. 扼要说明前驱的古典及几何定义与 Колмогоров 公理系的"联系"（同时将看到从口袋抽球是独立事件的概型）；3. 以 Колмогоров 公理系为基础，叙述初学必读的内容（同时将看到当从实变函数论角度理解概率收敛等概念时丝毫没有涉及概率内容）；4. 尽可能指出使用的数学工具之多种多样． 第一章到第三章 § 3.1 例 3.1.6 止只需要很少一点微积分知识，可作一般读物用．需要专业知识的只是其余的部分．因作者水平所限，书中疵谬难免，希望读者惠予赐教．

关肇直先生在百忙中审阅了本书稿，提出了许多宝贵意见；王寿仁、陈荫枋、程京、宫学惠、崔士英、王梓坤等诸先生分别在概率论、统计、物理、微分方程各方面给了作者很大的帮助，作者在此致以深切的谢意．

杨宗磐识

一九六〇年春

目　　次

第一章　概率论的古典构成

§ 1.1　概率的古典定义

用数学处理自然现象的时候，首先须将现象"概型化"．只有认为所有可能的状态，用一定的数学工具可以完全描述(有必要的时候，引进一定数目的参数)的时候，才能利用数学去考虑．这样一个数学上明确的系统，并不是实际本身，而是由实际中抽象出来的一个概型（Схема, Schema)，它深刻地反映实际，因之可供描述实际．

在古典力学里，只使用这样的概型，即在时刻 t 的状态 y，由时刻 t_0 的状态 x 一意确定；其数学表示为

$$y = f(x, t_0, t).$$

这样的一意函数 f 存在的时候，如在古典力学中经常假设的，则称这概型是一个"决定论"过程的概型． 有时候，状态 y 不完全由时刻 t_0 的状态确定，而与时刻 t 以前的变动有关． 不过通常这种依属关系可以避免，只须引进新的参数，扩大时刻 t 的系统的状态就行．

例如，质点的运动由运动方程所决定：

$$m\frac{d^2x_i}{dt^2} = X_i,\ i = 1, 2, 3,$$

这里 m 是质量，(x_1, x_2, x_3) 是质点的位置，(X_1, X_2, X_3) 是力的分量分解，X_i 是 (x_1, x_2, x_3)，$\left(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \frac{dx_3}{dt}\right)$ 及 t 的函数． 解出来，得到

$$x_i = x_i\left(t;\ x_1^0, x_2^0, x_3^0;\ \frac{dx_1}{dt}\Big|_{t_0}, \frac{dx_2}{dt}\Big|_{t_0}, \frac{dx_3}{dt}\Big|_{t_0}\right),\ i = 1, 2, 3.$$

在古典力学里，当解不能唯一确定时，通常理解为运动法则本身不

成立，也就是不够“力”的资格.

譬如，发射炮弹的时候，大气情况，炮身倾斜，炮弹初速确定了之后，炮弹的路线就完全确定了. 这就是最简单、最普遍的决定论概型:

I. 只要一组条件 $\mathfrak{S}$ 一实现，则事件 A 必定发生. 这样的事件叫作必然事件. 通常的大气情况、炮身倾斜、炮弹初速 ($\mathfrak{S}$) 已确定时，炮弹的路线是与地面平行的直线 (A)，是不会发生的. 这样的事件叫作不可能事件.

进一步看，局部的大气情况，受更大范围的大气情况的影响，变化是复杂的；炮身倾斜，因射击时的振动，难免有些摇摆，射击的目标，通常不大，那末，对击中目标来说，这些因素就不容忽视. 通常所说的大气情况、炮身倾斜、炮弹初速只是一种“平均”的性质. 因之条件组 $\mathfrak{S}$ 实现的时候，“击中目标”这一事件 A 就可以发生，可以不发生. 这样的 A 叫作随机事件.

随机事件的例还很多. 譬如，将钨丝保持在一定的高温，测定它所产生的热电流. 物理学家可以采取两种办法：一种是取很多份条件都一样的钨丝来测定；一种是只取一份，在长时间内保持一定条件来测定(为什么这两种作法结果可以一致，当然另有理由). 采取第一种办法的时候，是否每份的测定值都一样，却难说；采取第二种办法的时候，是否每个时刻的测定值都一样，也难说. 作实验的条件(钨丝的大小，周围的气压温度等宏观条件)尽管一样，个别的测定值，却是随机事件.

在发射炮弹的时候，可以重复多次发射，例如看 1000 次发射里击中目标(这个概念当然也要确切地定义)有若干次. 在测定钨丝的时候，当然可以看取某个值的份数有多少，或者看取某个值有多少次. 对于大量重复现象、集团现象和长时间现象，就建立了这样一个概型:

II. 在条件组 $\mathfrak{S}$ 实现之下，事件 A 发生的概率是 p. 这概率 p 反映所研究的现象的某些客观规律. 概率论就是研究概型II的规律性的数理科学. 也就因为这个原因，不但研究自然现象、工程技

术需要概率论，在社会生活的各方面也大量地应用着.

引进若干必要的名词.

1）蕴涵　使事件A发生的条件组$\mathfrak{S}$的每次实现之下，事件B也发生的时候，称A蕴涵B，记作$A\subset B$．例如$\mathfrak{S}$是大气压力78cm，加热到100℃，A是100℃的水，B是50℃以上的水.

2）等价　条件组$\mathfrak{S}$每次实现之下，A蕴涵B，B也蕴涵A，即A，B同时发生或不发生的时候，称A，B是等价的．记作$A=B$．所有必然事件都是彼此等价的，用U表示；不可能事件亦然，用$\varnothing$或V表示．任何事件蕴涵U，$\varnothing$蕴涵任何事件.

3）积　由A，B组成一个事件"A与B同时发生"，叫作A与B的积，记作$A\cap B$或AB．例如$\mathfrak{S}$是完全均匀的两个骰子，投掷一次．事件A是其中至少一个出现奇数，B是其中至少一个出现偶数．$A\cap B$等价于"出现的两数之和是奇数".

4）和　由A，B组成一个事件"A，B至少发生其中一个"，叫作A与B的和，记作$A\cup B$或$A+B$．例如取3）中的A，B，则$A\cup B$是两个骰子中至少一个的点数是奇数(两个同时是奇数的情形包括在内)，或者两个骰子中至少一个的点数是偶数(两个同时是偶数的情形包括在内)．显然$A\cup B$等价于"两个骰子点数之和$\geqq 2$".

注　和，积两概念可以推广到任意有穷个事件上去.

5）差　由A，B组成一个事件"A发生而B不发生"，叫作A与B之差．记作$A\backslash B$或$A-B$．例如，取3）中的A，B，则$A\backslash B$等价于"出现两个奇数".

6）对立事件　两事件A，A^c如果满足$A\cup A^c=U$，$A\cap A^c=\varnothing$的时候，称$A^c(A)$是$A(A^c)$的对立事件．例如3）中A的对立事件A^c是"同时出现两个偶数".

7）互不相容　A，B不能同时发生，即$A\cap B=\varnothing$时，称A，B是互不相容的.

8）　$A=B_1\cup B_2\cup\cdots\cup B_n$，$B_i\cap B_j=\varnothing(i\neq j)$

的时候，说将A划分成基本事件B_1，B_2，$\cdots$，B_n．例如掷骰子一

次 ($\mathfrak{S}$), A 是出现奇数, B_1, B_2, B_3 分别是出现 1 点, 3 点, 5 点.

9) 事件体 $\mathfrak{S}$ 实现之下,某些随机事件的组 $\boldsymbol{S}$ 若满足: 事件 A, B 属于 $\boldsymbol{S}$(用 $A, B \in \boldsymbol{S}$ 表示),则 $A \cup B$, $A \cap B$, $A \backslash B \in \boldsymbol{S}$, 又 U, $\varnothing \in \boldsymbol{S}$, 则称 $\boldsymbol{S}$ 为事件体.

以上引进的名词及事件、基本事件,在演绎系统里是没有方法再定义了. 基本事件是组成事件的"单位",任何事件认为都是可以划分成基本事件的. 基本事件随研究问题的不同,在同一对象上,可以有不同的划分. 例如,在掷骰子,注意的是出几点,基本事件是

出 1 点,出 2 点,出 3 点,出 4 点,出 5 点,出 6 点;注意的是出偶数点,出奇数点的时候,基本事件可以是出 1, 3,5 点及出 2, 4, 6 点.

例 1.1.1 $A \cap B \cap C = A$ 表示什么意义?

$A \cap B \cap C$ 表示 A, B, C 三事件同时发生. $A \cap B \cap C = A$ 表示 A, B, C 同时发生与 A 发生等价. 因之, A 发生的时候, B, C 也同时发生, 即 A 蕴涵 B, A 蕴涵 C: $A \subset B$, $A \subset C$. 反之, 设 $A \subset B$, $A \subset C$. 于是, A 发生的时候, A, B, C 同时发生, 即 $A \subset A \cap B \cap C$. A, B, C 同时发生, 则蕴涵 A 是显然的(即与 $A \subset B$, $A \subset C$ 无关). 换句话说, $A \cap B \cap C = A$ 与 $A \subset B$, $A \subset C$ 等价.

例 1.1.2 化简 $(A \cup B) \cap (B \cup C)$.

$A \cup B$ 表示或 A 发生,或 B 发生. $B \cup C$ 表示或 B 发生或 C 发生. $(A \cup B) \cap (B \cup C)$ 表示 $A \cup B$, $B \cup C$ 同时发生, 因之, 分四种情形:

(i) A 发生, B 发生,这时候蕴涵 B 发生;

(ii) A 发生, C 发生即 $A \cap C$ 发生;

(iii) B 发生, B 发生即 B 发生;

(iv) B 发生, C 发生,这时候蕴涵 B 发生.

所以 $(A \cap B) \cap (B \cup C)$ 蕴涵 $A \cap C$ 或 B 发生: $(A \cup B) \cap (B \cup C) \subset (A \cap C) \cup B$.

反之,设 $A \cap C$ 或 B 发生. 分两种情形:

(i) B 发生,则蕴涵 $A\cup B$, $B\cup C$ 同时发生;

(ii) $A\cap C$ 发生也蕴涵 $A\cup B$, $B\cup C$ 同时发生.

所以 $(A\cap C)\cup B$ 蕴涵 $A\cup B$, $B\cup C$ 同时发生:

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

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

就是欲求的化简.

例 1.1.3 证 $(A^c\cap B^c)^c=A\cup B$.

A^c, B^c 分别表示 A 不发生, B 不发生. $(A^c\cap B^c)^c$ 表示 A, B 同时不发生的对立事件,即或 A 或 B 要发生. 即 $(A^c\cap B^c)^c\subset A\cup B$. $A\cup B$ 发生,即或 A 或 B 发生. 分两种情形:

(i) A 发生, A^c 自不能发生, $A^c\cap B^c$ 更不能发生;

(ii) B 发生, B^c 自不能发生, $A^c\cap B^c$ 更不能发生.

所以 $A\cup B\subset(A^c\cap B^c)^c$. 合并得到的两个蕴涵,得所欲证.

随机事件的概率,用统计的方法求出的情形很多,也有先理论地考虑,再由事实验证的. 不管那一种情形,两种先验的想法是非常重要的. 在这一节,先谈第一种,所谓**概率的古典定义**.

概率的古典定义是把概率概念化为事件的等可能性. 例如掷骰子,它的形状是很准确的正立方体,并且材料是完全均匀的,掷向的平面是完全均匀的,由于骰子的对称性,任何一面客观上没有比其余的更特别的地方. 其定义如下:

事件 A 可划分为 m 个基本事件, 而这些基本事件都属于由 n 个两两互不相容并且等可能的基本事件所构成的事件体, 即事件体的必然事件可划分为这 n 个基本事件. 则 A 的概率 $P(A)=\frac{m}{n}$. 组成 A 的基本事件,又叫作**有利于 A 的基本事件**,所以上述定义又可以写成

$$A\text{ 的概率 }P(A)=\frac{\text{有利于 }A\text{ 的基本事件数目}}{\text{基本事件总数}}.$$

事件概率有下列三个基本性质:

1° $A\in\boldsymbol{S}$, 则 $P(A)\geqq0$;

2° $P(U)=1$;

3° **加法定理** $A=B\cup C$，$B\cap C=\varnothing$，$A,B,C\in \boldsymbol{S}$的时候，则 $P(A)=P(B)+P(C)$.

1°—3° 的证明及由 1°—3° 推出的若干结果，均从略（见 Гнеденко [1]，§ 4）. 进而陈述

乘法定理 两事件 E_1, E_2 其中一个的发生与否不影响另一个发生的时候，称 E_1, E_2 独立. 于是，$P(E_1\cap E_2)=P(E_1)\cdot P(E_2)$.

证 将 E_1, E_2 发生及不发生的可能情形，分成下列四种：

(i) E_1, E_2 同时发生，其发生数为 a,

(ii) E_1 发生而 E_2 不发生，其发生数为 b,

(iii) E_1 不发生而 E_2 发生，其发生数为 c,

(iv) E_1, E_2 同时不发生，其发生数为 d.

$$a+b+c+d=N.\quad P(E_1\cap E_2)=\frac{a}{N}.$$

E_1 发生的总数是 $a+b$，所以 $P(E_1)=\dfrac{a+b}{N}$. E_2 发生的总数是 $a+c$，所以 $P(E_2)=\dfrac{a+c}{N}$. 因为 E_1, E_2 是独立的，$P(E_2)$ 与 E_1 发生时 E_2 发生的概率是相等的，后者等于 $\dfrac{a}{a+b}$.

$$P(E_1\cap E_2)=\frac{a}{N}=\frac{a+b}{N}\cdot\frac{a}{a+b}=P(E_1)\cdot P(E_2).$$

证完.

注 独立也是一个非常基本的概念. 这个概念及定理可以推广到任意个“总起来独立”的事件. 详见§2.2. 在这一章，暂时承认这定理对有穷个事件的推广.

例 1.1.4 事件 A 的出现概率是 p，在 n 次独立试验中至少出现一次的概率是多少？

$1-p$ 是 A 不出现的概率，在 n 次独立试验，A 一直没有出现的概率，根据乘法定理是 $(1-p)^n$. 所以欲求的概率是 $1-(1-p)^n$.

进一步看，欲使 $1-(1-p)^n\geqq\dfrac{1}{2}$，$n$ 多大才行？不难算出