概率论引论

汪仁官

北京大学出版社

概率论引论

汪仁官

北京大学出版社

新登字(京)159号

书　　名：概率论引论
著作责任者：汪仁官
责任编辑：王明舟
标准书号：ISBN 7-301-02558-0/O·340
出 版 者：北京大学出版社
地　　址：北京市海淀区中关村北京大学校内　100871
电　　话：出版部 2502015　发行部 2559712　编辑部 2502032
排 印 者：北京大学印刷厂
发 行 者：北京大学出版社
经 销 者：新华书店
版本记录：850×1168毫米　32开本　9.125印张　230千字
1994年4月第一版　1994年4月第一次印刷
印　　数：0001—5,000册
定　　价：6.50元

序　言

概念论是研究自然界、人类社会及技术过程中大量随机现象中的规律性的一门数学。随着现代科学技术的迅速发展和人类生活条件的不断改变，这门数学学科得到蓬勃的发展。它不仅形成了系统的理论，而且在自然科学、人文科学、工程技术及经营管理等方面有越来越多的应用。许多院校都开设概率论（或概率统计）课。

多年来，北京大学数学系和概率统计系都把概率论列为学生必修的基础课。为使学生尽早具有概率的直觉，认识概率论和实际事物的紧密联系并产生浓厚的学习兴趣，从八十年代初开始将“概率论”课安排在学生第二学年的第二学期，因而在“实变函数”课之前。怎样编写适合教学需要的概率论教材？有一个问题需要很好解决。就是如何处理逻辑严谨性与生动直觉的关系，使学生既有严谨的抽象思维的能力又有概率统计的直觉，对实际事物中的随机性产生敏感。众所周知，有了实变函数和测度论的知识，概率论课中的定理都可获得最一般情形下准确的叙述和严格的证明。但学完实变函数和测度论后再学概率论，对于四年制大学生来讲是太迟了。一方面，对于概率统计专业的学生来讲，很多后继课程难以安排；另一方面，容易给部分学生造成一种错觉，以为概率论不过是测度论的一种应用，不需要去钻研人类长期积累起来的丰富的概率思想，不去发展朴素的概率直觉，因而不利于培养学生学习与研究概率论的积极性。既然“概率论”课放在“实变函数”课之前，不可能在这门课里对所有叙述的定理都给出证明，如何保持课程的逻辑严谨性并有利于学生概率直觉能力的培养就是教材建设中必须处理好的重要问题。

汪仁官教授长期担任北大数学系和概率统计系“概率论”的讲授工作，积累了丰富的经验，曾获北京大学优秀教学奖与教学成果奖。这次他在自编的油印讲义的基础上进行了较大的扩充与改写，形成了本书。我认为，这本书取材适当、文字简练、通俗易懂，是适合教学需要的好书，有下列特点：

1. 注重培养学生对随机现象的理解和概率直觉。例如在介绍现代公认的 kolmogorov 概率公理化定义之前还全面介绍概率的其它几种定义（古典定义、几何定义 统计定义），使学生了解人类形成的丰富的概率思想，更好地理解概率的公理化定义。为了使学生了解随机现象的广泛性和丰富多采并理解独立性、相依性、条件概率等重要概念，不仅列举了大量实际例子，而且设专章介绍齐次泊松过程和马尔可夫链这些具体的生动的随机过程，而不去对随机过程的一般概念进行抽象讨论。

2. 坚持数学上的严谨性，培养学生严密思维的能力。凡是利用数学分析课和高等代数课的内容可以证明的定理都给出严格的证明；凡是要用到实变函数和测度论才能给出完全证明的结论，则明确声明本书不给出证明。但所述结论的实际意义仍加以阐述，并尽可能在一些特殊条件下加以论证。例如，多个随机变量的函数的均值公式就介绍而不给出一般性证明；对特征函数与分布函数的关系只叙述结论而不加以证明；对中心极限定理只在一些特殊情形下利用特征函数加以证明；等等。

我相信，本书的出版将对我国的概率论教学很有益处。

陈家鼎

北京大学概率统计系

一九九三年十二月

目　录

第一章　古典概型与概率测度的公理化

§1　古典概型

所谓某事件在一次试验（或一次观察）中发生的概率，是试图用一数量来刻划该事件在一次试验中发生的可能性的大小。在某些问题中，根据问题本身具有的"对称性"，充分利用人类长期积累的关于"对称性"的实际经验，分析事件的本质，是可以合理地规定与找出事件的概率来的。

1．古典概型中对概率的定义

定义　若全体试验结果是 n 个，且是等可能的；而有利于事件 A 的试验结果有 m 个，则规定 m/n 为事件 A 的概率．记作

$$P(A)=m/n. \tag{1.1}$$

2．几个例子

例1.1　投掷一枚分币．规定 $A=$"国徽朝上"．求 $P(A)$。

解　该试验有两个结果："国徽朝上"、"国徽朝下"，它们是等可能的（假定该分币是匀称的）．由(1.1)，

$$P(A)=1/2.$$

例1.2　投掷一颗骰子．规定 $A=$"四点"，$B=$"偶数点"．求 $P(A),P(B)$。

解　该试验有六个结果："一点"、"二点"、…、"六点"，它们是等可能的（假定该骰子是匀称的）．显然，

$$P(A)=1/6.$$

对于 B，有利于它的试验结果为“二点”、“四点”、“六点”，因此 $m=3$。于是由(1.1)，

$$P(B)=3/6=0.5.$$

例1.3 投掷两枚分币，求国徽“一上一下”的概率。

解 该试验有四个等可能的结果：

$$A_1=\text{“上、上”},\qquad A_2=\text{“上、下”},$$
$$A_3=\text{“下、上”},\qquad A_4=\text{“下、下”}.$$

由(1.1)，B(=“一上一下”)的概率为：

$$P(B)=2/4=0.5.$$

例1.4 一盒中有五个球(三白二红)，现从中随机取两个。求两个都是白球的概率。

解 给五个球编号如下：

该试验有 C_5^2 个结果，它们是等可能的(这是该问题中“随机”二字的确切含意)。而有利于事件“两个都是白球”的试验结果共 C_3^2 个。由(1.1)，若记 A=“两个都是白球”，则

$$P(A)=C_3^2/C_5^2=3/10.$$

§2 数 数

在古典概型中，为找出事件 A 的概率，据(1.1)，只需数出两个数：一个是全体试验结果数 n，另一个是有利于 A 的试验结果数 m。为此，在本节中专门来讨论与此有关的计数的基本知识与技能。

1. 抽样模型

例2.1 盒中有 4 个白球，6 个红球，现从中随机取 4 个。求：

取到 2 个白 2 个红的概率。

解 用例 1.4 的考虑方法，有 $n=C_{10}^4$；而 $m=C_4^2\cdot C_6^2$（为什么？），因此

$$P(2白2红)=C_4^2C_6^2/C_{10}^4=3/7.$$

有人按另一考虑得 $P(2白2红)=A_4^2A_6^2/A_{10}^4$。这结果正确否？问题在哪里？

推广至一般，我们有如下结果。

设一批产品共 N 个，其中有 M 个次品（其余的是正品），现从中随机取 n 个。则恰好取到 m 个次品的概率

$$P(A)=C_M^m\cdot C_{N-M}^{n-m}/C_N^n. \tag{2.1}$$

更一般地，我们有：

设有 N 个东西，其中第 i 类有 N_i 个东西（$i=1,\cdots,k$），$N_1+\cdots+N_k=N$，现从中随机取 n 个。则事件 $A=$"恰好有 n_i 个属于第 i 类，$i=1,\cdots,k$"的概率

$$P(A)=C_{N_1}^{n_1}\cdot C_{N_2}^{n_2}\cdots C_{N_k}^{n_k}/C_N^n, \tag{2.2}$$

其中 $\sum\limits_1^k n_i=n$。

显然，(2.1)是(2.2)的特例。

练习题： 桥牌比赛中，南家拿到 4 张黑桃，且其它花色都是 3 张的概率是多少？

2. 二项系数与多项系数

我们知道组合数 C_n^m 是这样一个数：

从 n 个不同的元素中，取 m 个，不可重复，不管次序，问共有多少种不同的结果？我们将这结果数记为 C_n^m。

由此可知，二项式 $(1+x)^n$ 展开式中 x^m 的系数应是 C_n^m（为

什么？你能直接按 C_n^m 的上述定义来说明吗？)。因此我们也称组合数 C_n^m 为二项系数。

我们还知，将 n 个不同元素，分为有次序的两个组，它们分别有 $m,n-m$ 个元素，共有 C_n^m 种不同的分法。那么，将 n 个不同元素，分为有次序的 r 个组，各组分别有 $n_1,\cdots,n_r(\sum n_i=n)$ 个元素，共有多少种分法？

这个问题是前一问题的推广，而且可用前一问题的结论来解决后一问题。因为可将分 r 个有序组的问题分解为分 2 个有序组的问题：先分为$n_1,n-n_1$ 两个有序组，再将 $n-n_1$ 分为 $n_2,n-n_1-n_2$ 两个有序组，…，因此，共有

$$C_n^{n_1}C_{n-n_1}^{n_2}\cdots C_{n-n_1-\cdots-n_{r-1}}^{n_r}=\frac{n!}{n_1!n_2!\cdots n_r!}\qquad(2.3)$$

种分法。

显然多项式$(x_1+\cdots+x_r)^n$ 的展开式中，$x_1^{n_1}\cdots x_r^{n_r}$ 的系数应是(2.3)(为什么？)。因此我们称(2.3)右边那个数为多项系数。

练习题：桥牌比赛中，出现某特定牌（比如，东家拿到全部黑桃，南家拿到全部红心，…）的概率为多少？

3. 可重复组合数

(1) 关于 n 维 0-1 向量

所谓 n 维 0-1 向量，是分量仅取 0 或 1 的 n 维向量。有两个结论大家是熟悉的：

a. n 维 0-1 向量共有 2^n 个；

b. n 维 0-1 向量恰含 m 个 0 的向量数是 C_n^m 个。

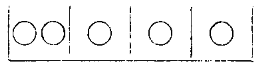
(2) 关于可重复组合数

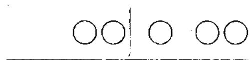
从 n 种不同的元素中取 m 个元素有多少个不同的结果？按是否可以重复及是否(只管内容)不管次序，有下列情况：

	不可重复	可重复
管 次 序	A_n^m	n^m
不管次序	C_n^m	?

为找出所谓可重复组合数，我们具体分析一下 $n=4, m=5$ 的情形。利用所谓球匣模型来求解是颇有趣味的。

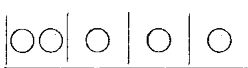
考虑 4 只有序的匣子，共装有 5 个不可分辨的球。不难看出，一个可重复组合对应到一种球在匣中的分配情况，例如：

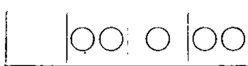
$AABCD$ $\longleftrightarrow$ 

$BBCDD$ $\longleftrightarrow$ 

再注意到，这种对应在全体可重复组合与全体球在匣中的分配情形是一一对应的。因此为数出可重复组合(总)数，只需数出球的分配情况(总)数。

那么，球的分配情况有多少种呢？我们再把球的分配情况与 0-1向量对应起来：

 $\longleftrightarrow$ (0 0 1 0 1 0 1 0)

 $\longleftrightarrow$ (1 0 0 1 0 1 0 0)

(球对应到“0”,什么对应到“1”呢？)

注意，这种 0-1 向量是 8 维的，且恰有 5 个0；更重要的是，这对应在全体球的分配情况与全体这种 0-1 向量间是一一对应的。至此，得到问题的答案：从 4 种不同的元素中取 5 个元素的可重复组合数是 C_8^5。

从以上的分析，可直接得一般结论：

(从 n 中取 m 的)可重复组合数为

$$C^{m}_{n+m-1}. \qquad (2.4)$$

练习题：$(x+y+z)^5$ 的展开式中，有多少种不同的项？

4. 怎样求解

怎样利用有关数数的基本公式，来正确（不漏、不重）数出具体问题中的数，并不是一件轻而易举的事。这里我们引用钟凯莱教授在《初等概率论附随机过程》一书中提出的几条忠告，也许是有益的。

（1）如果你看不清你的问题，用小的数目试一个特殊（但不要太特殊）的情况使你能看得较清楚一些。这将使你在思想上明确，要数的是什么；而且，特别是在发现重复和遗漏两方面会对你有所帮助。

（2）把问题分解成若干部分，如果这些部分较简单、较清楚并且较容易集中你的思考的话，这有时可以通过固定许多"变量"中的一个来实现。

（3）如果你见到复杂性升高得很快，不要试图逐步推敲。在我给学生所有的劝告中，这一条最不受重视；然而可能最有收获。对前两步，一步一步的计数看来可能是容易的，但是你能看到怎样把它进行到底吗？

（4）如果问题的叙述中有模棱两可的地方，不要回避。这是一个语义上的问题。如有必要，试一试各种解释。无论如何，不要利用文字的不确定性或者教师的疏忽，而把一个合理的问题变成一个没有什么意义的问题。

下面我们再看几个例子。

例2.2 把一副扑克牌① 洗透，四张 A 连在一起的概率是多少？

① 本书中，一副扑克牌均指已去掉大小王牌，即全副牌共52张。

解 一副扑克牌共52张，因此共有52!个排列。所谓"洗透"是指各排列是等可能的。按古典概型(1.1)，$n=52!$。为求m，即要求出多少种情况是4张A连在一起的，这可先考虑4张A在整副牌的前4张的情况(这里，实际上是在按忠告(2)行事)，共有$4!\cdot 48!$种。至于4张A的位置，由于它们是连在一起的，应有49种(为什么?)。综上所述，所求概率为

$$\frac{49\times 4!\times 48!}{52!}=\frac{24}{52\times 51\times 50}=1.81\times 10^{-4}.$$

例2.3 掷六颗骰子，得到三对的概率是多少?

解 我们首先集中注意于计算"一对1，一对2，一对3"的情况数(这还是在按忠告(2)行事)。这可考虑为将6个不同的元素，分为有次序的3个组，每组2个元素，按(2.3)，共有分法$6!/2!2!2!$种。

至于哪三对? 有多少种? 显然是C_6^3。于是由(1.1)，

$$P(\text{"三对"})=C_6^3(6!/2^3)/6^6=25/648=0.0386.$$

练习题：掷六颗骰子，得到一对的概率，二对的概率分别为多少? 与你的印象大致相符吗?

例2.4（生日问题） n人中至少两人有相同生日的概率是多少?

解 它的"对立事件"——n人中没有两人有相同生日——的概率是容易求得的。因此，所求的概率是①

$$P_n=1-A_{365}^n/365^n.$$

经计算得下表

① 由§5中(5.1)知，两个对立事件的概率之和为1.

n	P_n
5	0.03
10	0.12
15	0.25
20	0.41
21	0.44
22	0.48
23	0.51
24	0.54
25	0.57
30	0.71
35	0.81
40	0.89
45	0.94
50	0.97
55	0.99

出人意外的是当 $n\geqslant 23$ 时，这一概率就大于0.5。你们班有多少人，有相同生日的吗？

5．一个著名问题——匹配问题

四张卡片分别标着1,2,3,4,面朝下放在桌子上。一个自称有透视能力的人将用他超感觉能力说出卡片上的号数，如果他是冒充者而只是随机地猜一下，他至少猜中一个的概率是多少？

对于这个小数目($n=4$)的具体问题，可以通过把“至少猜中一个”进行分析而获得解答。这里仅给出分析结果：

恰好猜中的个数	情况数	概率
4	1	1/24
3	0	0
2	6	1/4
1	8	1/3
0	9	3/8

可是当 n 较大时，比如 $n=50$，怎么办？这推动人们去建立（或寻找）一些基本的公式(见§5)。

§3 几何概型

在古典概型中利用等可能性，成功地计算了某一类问题的概率；不过，古典概型要求可能结果的总数为有限。对于有无限多可能结果而又具有某种等可能性的场合，一般可用几何方法来求解。

例3.1（约会问题） 两人相约某天5点至6点在某地会面。先到者等候另一人20分钟，过时就离去。试求这两人能会面的概率。

解 以 x,y 分别表示两人到达的时刻，则会面的充要条件为 $|x-y|\leqslant 20$。记

$$G=\{(x,y)\mid 0\leqslant x\leqslant 60, 0\leqslant y\leqslant 60\},$$

$$g=\{(x,y)\mid |x-y|\leqslant 20,(x,y)\in G\}.$$

两人能会面的概率(参见图 1-1)

$$P(A)^{(*)}=\frac{g\text{ 的面积}}{G\text{ 的面积}}=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

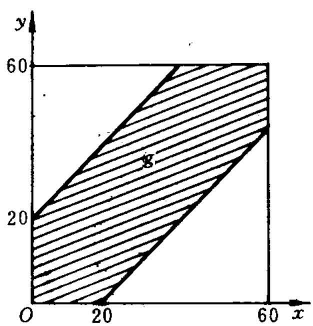


图 1-1

（*）这个等号，实际上是定义；是在两人到达的时刻在5点至6点范围内机会均等，互不影响的背景下作出的规定。

我们也可写出类似于古典概型（1.1）式的一般公式：若全体试验结果的图形表示为G，且机会均等；而有利于事件A的试验结果的图形表示为g，则规定

$$P(A)=\frac{S(g)}{S(G)}, \tag{3.1}$$

这里的$S(\cdot)$,可以代表图形的面积,也可以是长度、体积等。

例3.2（蒲丰（Buffon）问题） 平面上画着一些平行线，它们之间的距离都等于a,向此平面任投一长度为$l(l<a)$的针。试求此针与某一平行线相交的概率。

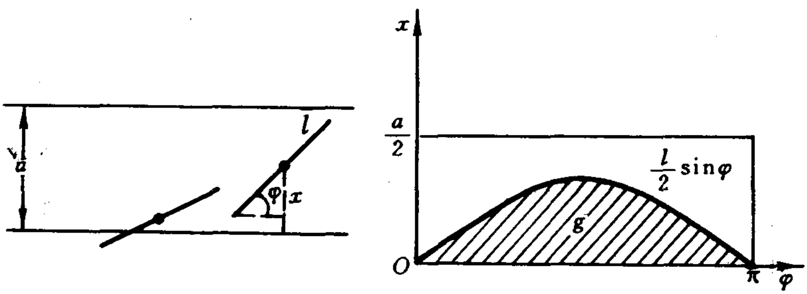


图 1-2

解 以x表示针的中点到最近的一条平行线的距离,φ表示该平行线与针的交角(见图1-2)。显然有

$$0\leqslant x\leqslant\frac{a}{2},\quad 0\leqslant\varphi\leqslant\pi,$$

因此

$$G=\left\{(\varphi,x)\,|\,0\leqslant\varphi\leqslant\pi,\quad 0\leqslant x\leqslant\frac{a}{2}\right\}.$$