广播电视大学教材

概率统计讲义

陈家鼎 刘婉如 汪仁官 编

人民教育出版社

内容提要

本书介绍了概率统计的基本内容和某些应用，叙述清楚、简明易懂，要求读者具有普通的微积分和一点线性代数知识．可供广播电视大学学生使用．

广播电视大学教材

概率统计讲义

陈家鼎 刘婉如 汪仁官 编

*

人民教育出版社出版

新华书店北京发行所发行

北京印刷一厂印装

*

开本 787×1092 1/32 印张 10 字数 225,000

1980 年 7 月第 1 版 1981 年 3 月第 2 次印刷

印数 191,501—342,500

书号 13012·0486 定价 0.73 元

序　言

革命导师恩格斯说："在表面上是偶然性在起作用的地方，这种偶然性始终是受内部的隐蔽着的规律支配的，而问题只是在于发现这些规律."（见《马克思恩格斯选集》第四卷，第243页，1972年版）. 偶然事件的概率（即发生的可能性的大小）就是该偶然事件隐蔽着的特性. 概率论与数理统计就是研究这种内在特性的一门数学学科. 随着现代科学技术的迅速发展，这门数学学科也得到了蓬勃的发展. 它不仅形成了结构宏大的理论，而且在很多科学研究、工程技术和经济管理的领域里有愈来愈多的应用. 由于应用的广泛性，许多理工科专业（以及经济系科）都把"概率统计"列为学习课程，培养学生处理随机现象的能力.

去年8月，中央广播电视大学要北大数学系承担"概率统计"课的教学任务，加上校内一些系也需开设这门课，这促使我们考虑教材问题. 这本"概率统计讲义"就是在这种形势的推动下，在我们以前编写的同名讲义的基础上，经过较大的改写、扩充而成的.

在这次编写工作中，我们注意了下列几点：

(1) 本书是针对50～70学时的讲课需要而编写的，只能讲解概率统计的一些基本内容与某些实用范围较广的方法，不能求多求全. 但讲解详细，便于自学. 凡小字排印部分都可略去不讲. 在学时较紧的情况下，除前三章必须有足够时间教学外，其它各章都可以略去一部分.

(2) 努力贯彻理论联系实际的原则，对基本概念、重要公式

和定理的实际意义多加解释，多举各方面的例子，力求通俗易懂，便于读者把所学的内容和实际工作结合起来．考虑到回归分析方法与正交试验法应用广泛，所以把它们分列专章，供读者选学．

(3) 虽然概率统计的严密的深入的数学理论不能离开实变函数论与测度论，但在目前情形下作为非数学专业用的概率统计教材，应该尽量少用专门的数学知识．这本讲义只用到普通的微积分知识，正文里基本上不用线性代数知识，有些结论不给出严密的数学证明．

一般说来，作为“概率统计”课的教材，应该有一章的篇幅介绍随机过程的最基本知识．但这次编写时间太紧，加上考虑到当前这门课教学时数的限制，故本书未涉及随机过程内容．

在这次编写过程中，我们参考了许多概率统计书籍和教材，特别是在例题和习题的选配方面，吸取了它们中的不少材料．我们还得到中国科学院系统科学研究所研究员张里千同志的帮助．谨在此致谢．

由于我们水平有限，加上编写时间仓促，书中的缺点、错误一定不少，欢迎读者批评指正．

编　者

1980年1月于北京大学数学系概率统计教研室

参 考 书 目

[1] Б. В. Гнеденко, 概率论教程, 1956. (丁寿田译)

[2] M. Fisz, 概率论及数理统计, 1962. (王福保译)

[3] 赵仲哲, 概率论讲义, 1957. (北京大学内部油印)

[4] 林少宫, 基础概率与数理统计, 1963.

[5] 沈恒范, 概率论讲义, 1966.

[6] 王梓坤, 概率论基础及其应用, 1976.

[7] 中国科学院数学研究所概率统计室, 回归分析方法, 1974.

[8] 中国科学院数学研究所概率统计室, 常用数理统计表, 1974.

[9] 北京大学数学系试验设计组, "电视讲座: 正交试验法", 1979.

[10] 北京大学数学力学系数学专业概率统计组, 正交设计, 1976.

[11] 浙江大学数学系高等数学教研组, 概率论与数理统计, 1979.

目 录

第一章 随机事件与概率

§1 随机事件及其概率

粗略地说，在一定的条件下，可能发生也可能不发生的事件，称为**随机事件**（更确切的叙述见下面的定义）.

例 1.1 投掷一枚分币，“正面朝上”这个事件（记作 A），是一个随机事件．在该试验中，“正面朝下”（记作 B），也是随机事件．（我们常把有国徽的一面称为正面．）

例 1.2 投掷两枚分币，则

A＝“两个都是正面朝上” $P(A)=\frac{1}{4}$

B＝“两个都是正面朝下”

C＝“一个正面朝上，一个正面朝下”

都是随机事件．不难看出

D＝“至少有一个正面朝上”

也是随机事件.

例 1.3 从十个同类产品（其中有 8 个正品，2 个次品）中，任意抽取三个．那么，

A＝“三个都是正品”

B＝“至少一个是次品”

均为随机事件．而

“三个都是次品”和“至少一个是正品”

这两个事件呢，前者是不可能发生的；后者是必定要发生的．我们称不可能发生的事件为**不可能事件**，记作 V；称必定要发生的

事件为**必然事件**，记作U. 为讨论问题方便起见，将不可能事件V和必然事件U也当作随机事件.

对于随机事件，在一次试验中是否发生，我们虽然不能预先知道，但是它们在一次试验中发生的可能性是有大小之分的. 比如，在例1.1中，如果投掷的分币是匀称的，那么，随机事件A(=“正面朝上”)和随机事件B(=“正面朝下”)发生的可能性是一样的；在例1.2中，如果两个分币都是匀称的，那么随机事件A(=“两个都是正面朝上”)和随机事件B(=“两个都是正面朝下”)发生的可能性也是一样的，并且它们比随机事件C(=“一个朝上，一个朝下”)发生的可能性要小. 不仅如此，由我们的直觉还可以说，发生例1.1中随机事件A(=投掷一枚分币出现“正面朝上”)的可能性，比发生例1.2中随机事件A(=投掷两枚分币，“两个都是正面朝上”)的可能性要大. 然而，对事件发生的可能性只停留在基本上是定性的了解与描述上，实在太不够了. 我们希望对它给出客观的定量的描述.

回到例1.1中投掷一枚分币的试验，这种试验是在一定条件下作的. 比如说，我们规定：“分币是匀称的，放在手心上，用一定的动作向上抛，让分币自由落在具有弹性的桌面上，等等.”称这些条件为条件组S. 于是，在条件组S的一次实现下，事件A(“正面朝上”)是否发生是不确定的. 然而这只是问题的一方面. 当条件组S大量重复实现时，事件A发生的次数，也称为**频数**，能体现出一定的规律性，约占总试验次数的一半. 这也可以写成

$$A\text{ 发生的}\textbf{频率}=\frac{\text{频数}}{\text{试验次数}},\text{接近于}\frac{1}{2}$$

在我们的心目中，由长期经验积累所得的，所谓某事件发生的可

能性的大小，不就是这个“频率的稳定值”吗？

历史上，有些人作过成千上万次投掷钱币的试验．下表列出他们的试验记录：

实验者	投掷次数 n	出现“正面朝上”的次数 μ（即频数）	频率 $=\mu/n$
DeMorgan	2048	1061	0.518
Buffon	4040	2048	0.5069
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005

容易看出，投掷次数越多，频率越接近 0.5.

定义① 在不变的一组条件 S 下，重复作 n 次试验．记 μ 是 n 次试验中事件 A 发生的次数．当试验的次数 n 很大时，如果频率 μ/n 稳定地在某一数值 p 的附近摆动；而且一般说来随着试验次数的增多，这种摆动的幅度愈变愈小，则称 A 为随机事件，并称数值 p 为随机事件 A 在条件组 S 下发生的概率，记作

$$P(A)=p$$

显然，数值 p 就成为 A 在 S 下发生的可能性大小的数量刻划．例如 0.5 就成为掷一枚分币出现“正面朝上”的可能性的数量刻划.

上述定义也可简单地说成：“频率具有稳定性的事件叫做随机事件，频率的稳定值叫做该随机事件的概率．”

① 我们这里只给概率一个直观的朴素的描述，通常称为概率的统计定义．精确的数学定义，即所谓公理化定义，在本课程中将不叙述．至于概率 $P(A)$ 的实际计算法，定义本身也给出了一种近似求法，即作大量的试验，计算事件 A 发生的频率．虽然得到的是近似值，但我们相信读者不至于因为现实生活中某一数值的获得只是些近似值而感到不实在．事实上，我们周围许多量的测量完全是近似的，如长度的概念并不会因为每次实测数值都是近似值而建立不起来，也不会因为温度计读数都是近似值而怀疑起“温度”的客观存在性.

我们强调指出，人类的大量实践证明，在实际中遇到的事件一般都是随机事件，也就是说都是有确定的概率的. 以后我们常简称随机事件为事件.

由于频率$\frac{\mu}{n}$总介于0,1之间，因而由概率的定义知，对任何随机事件A，有

$$0 \leqslant P(A) \leqslant 1$$

而对必然事件U及不可能事件V，显然有

$$P(U)=1, \quad P(V)=0$$

§2 古典概型

上面介绍了概率的定义. 它既是概念，同时又提供了近似计算概率的一般方法. 但是在某些特殊情况下，并不需要临时做多次试验，也就是说临时多次实现条件组S，从而求得概率的近似值，而是根据问题本身所具有的某种“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，就可以直接计算其概率.

例如上节的例1.1，即使我们不临时作大量的投掷试验，我们也会想到，“正面朝上”与“正面朝下”出现的机会相等. 因此，可以推测在大量试验中“正面朝上”这件事发生的频率在0.5左右，即它的概率为0.5. 为什么“正面朝上”与“正面朝下”机会均等呢？这是因为问题本身有一种对称性(匀称的分币)，如果“朝上”与“朝下”出现的机会不相等，那反倒与我们长期形成的“对称”的经验不相符了.

例2.1 盒中装有五个球(三个白球，二个黑球)从中任取一个，问：取到白球的概率是多少？

既然是“任取”，那末五个球被取到的机会一样，而白球有三个，因此，取到白球的概率应该是 3/5. 说得更清楚些，我们把五个球编上号如下(其中白球为 1, 2, 3 号；黑球为 4, 5 号)：

① ② ③ ④ ⑤

因为是随便取一个，所以

“取到 1 号球”，“取到 2 号球”，“取到 3 号球”

“取到 4 号球”，“取到 5 号球”

这些结果发生的机会一样，而且是互相排斥的，以及除此以外不可能有别的结果. 注意到 1, 2, 3 号球是白球，所以“取到白球”这个事件发生的频率会稳定在 3/5 左右，因此按概率定义，它的概率是 3/5.

例 2.2 盒中装有球的情况如上例，现从中任取两个，问两个球全是白球的概率是多少？

这个问题较为复杂，不过仍可按上例的方法进行分析. 还是把五个球同样编号，因为是随便取两个，所以下列这些结果

“①, ②”[①]，“①, ③”

“①, ④”，“①, ⑤”

“②, ③”，“②, ④”

“②, ⑤”，“③, ④”

“③, ⑤”，“④, ⑤”

发生的机会一样，而且是互相排斥以及除此之外不可能有别的结果，再注意到，上列十种情况中，有且仅有三种，即“①, ②”“①, ③”，“②, ③”为全白. 因此“全白”发生的频率会稳定在 3/10 左右. 于是，它的概率是 3/10.

---

① “①, ②”是“取到 1, 2 号球”的缩写. 下同.

推而广之，对上面几个例子所讨论的问题及解决问题的办法进行归纳，可得出一般规律.

定义 称一个事件组 $A_1, A_2, \cdots, A_n$ 为一个**等可能完备事件组**，如果它具有下列三条性质：

(1) $A_1, A_2, \cdots, A_n$ 发生的机会相同(等可能性)；

(2) 在任一次试验中，$A_1, A_2, \cdots, A_n$ 至少有一个发生.（也就是所谓"除此之外，不可能有别的结果".）(完全性)；

(3) 在任一次试验中，$A_1, A_2, \cdots, A_n$ 至多有一个发生.（也就是所谓"它们是互相排斥的".）(互不相容性).

等可能完备事件组在这里也称为**等概基本事件组**；其中任一事件 $A_i(i=1, 2, \cdots, n)$ 称为**基本事件**.

（在例 1.1 中，等概基本事件组的 $n=2$，它的两个基本事件是"正面朝上"与"正面朝下". 读者可对例 2.1 和例 2.2 分别找出等概基本事件组.）

若 $A_1, \cdots, A_n$ 是一个等概基本事件组，而事件 B 由其中的某 m 个基本事件所构成. 大量实践经验表明，事件 B 的概率应由下列公式来计算[①]：

$$P(B)=m/n \tag{2.1}$$

所谓古典概型就是利用关系(2.1)来讨论事件的概率的模型.

现在通过(2.1)式来讨论例 2.2. 考虑从三个白球两个黑

---

① 通常，如果试验只可能有有限个不同的试验结果 $A_1, A_2, \cdots, A_n$；而且它们发生的机会相同，则不难看出，$A_1, A_2, \cdots, A_n$ 就是一个等概基本事件组.（因此，解决这类问题主要是把 n 和 m 数出来.）

"只可能有有限个不同的试验结果"中的"试验结果"一词，是比较朴素的，直观的，方便的，一般而言，也是不会引起混淆的(今后我们有时也用这个词). 然而，毕竟不够准确，因此我们引进了等概完备事件组的概念.

球中任取两球，我们知道共有 $C_5^2=\frac{5\times4}{1\times2}=10$ 种不同的取法，它们出现的机会相同．每一种取法对应一个基本事件，所以等概基本事件组共含 $n=10$ 个事件（读者不难验证它的"完全性"和"互不相容性"）．而取得两球均为白球，共有 $m=C_3^2=\frac{3\times2}{2\times1}=3$ 种取法．由(2.1)有

$$P(\text{取得两个白球})=\frac{m}{n}=\frac{3}{10}$$

下面再看几个例子．

例2.3 设有一批产品共100件，其中有5件次品，现从中任取50件，问：无次品的概率是多少？ C_{95}^{50}/C_{100}^{50}

解 首先，从100件产品中任取50件，我们知道共有 C_{100}^{50} 个不同的结果，每一个结果就是一个事件．容易验证这些事件是一个等概基本事件组（是否等可能？是否完全？是否互不相容？读者自己想一想）．

现在来看 $B=$"任取50件其中无次品"，它由哪些基本事件所构成？多少个？很明显，要所取的50件中无次品，必须是从那95件正品中取来的．可见这种无次品的取法共有 C_{95}^{50} 种（即事件 B 含 C_{95}^{50} 个基本事件）．

由关系式(2.1)得

$$\begin{aligned}P(B)&=C_{95}^{50}/C_{100}^{50}\\&=\frac{95!/(50!45!)}{100!/(50!50!)}\\&=\frac{50\cdot49\cdot48\cdot47\cdot46}{100\cdot99\cdot98\cdot97\cdot96}\\&=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{47}{99}\cdot\frac{46}{97}\\&=\frac{1081}{38412}=2.8\%\end{aligned}$$

(请读者将本例跟例 2.2 进行比较.)

再考虑较为复杂的情形.

例 2.4 条件组 S 跟例 2.3 相同(即还是 100 件产品，其中有 5 件次品，从中任取 50 件)，问：恰有两件次品的概率是多少？

解 等概基本事件组同例 2.3，总数 $n=C_{100}^{50}$. 现在，问题的关键在于，计算出事件 $A=$“恰有两件次品”所包含的基本事件数.

取出的 50 件中，恰有两件次品，即有 48 件正品，2 件次品. 这 48 件正品必是从 95 件正品中取出的，共有 C_{95}^{48} 种；而 2 件次品必是从 5 件次品中取出的，共有 C_5^2 种. 因此，“恰有两件次品”共包含 $C_{95}^{48}\cdot C_5^2$ 个基本事件.

于是，据(2.1)得

$$
\begin{aligned}
P(A) &= C_{95}^{48}\cdot C_5^2/C_{100}^{50} \\
&= \frac{95!}{48!47!}\cdot\frac{5!}{2!3!}\bigg/\frac{100!}{50!50!} \\
&= 0.32
\end{aligned}
$$

就是说，任取 50 件，恰有两件次品的概率是 0.32.

习 题 一

1. 求例 1.2 及例 1.3 中的 $P(A)$, $P(B)$.

2. 设一批产品共 N 个，其中有 M 个次品(其它是正品)，现从中任取 n 个，问：恰有 m 个次品的概率是多少？($M<N$, $n<N$, $m\leqslant n$, $m\leqslant M$, $n-m\leqslant N-M$)

$C_{N-M}^{n-m}\cdot C_M^m/C_N^n$

3. 某人有 5 把钥匙，但忘记了开房门的是哪一把，逐把试开，问：(i) 恰好第三次打开房门锁的概率是多少？(ii) 三次内打开的概率是多少？(iii) 如 5 把内有 2 把房门钥匙，三次内打开的概率是多少？

$P_3^1\cdot P_4^4\quad P_5^5$

$P_4^4/P_5^5 = 1/5$

$(C_2^1P_3^1\cdot P_2^1\cdot P_3^3+P_3^2P_3^3)/P_5^5 = (72+36)/120 = 108/120 = 18/20 = 9/10$

4. 袋中有红、黄、白色球各一个，每次任取一个，有放回地抽三次，求下列事件的概率：

A＝"三个都是红的"＝"全红" $1/3^3$

B＝"全黄"，C＝"全白"，D＝"颜色全同" $3/3^3$ $2^3/3^3$

E＝"全不同"，F＝"不全同"，G＝"无红" $P_3^3/3^3=6/27=2/9$ $(3^3-3)/3^3=8/9$

H＝"无黄"，I＝"无白"，J＝"无红且无黄" $1/3^3$

K＝"无红或无黄"，L＝"全红或全黄" $2\cdot 1/3^3$

5. 从一付扑克的52张牌中，任意抽出两张，问都是黑桃的概率？ C_{13}^2/C_{52}^2

6. 在例2.4中，求至少有两件次品的概率.

$(2^3+2^3-1)/3^3=15/27=5/9$

事件的运算及概率的加法公式

我们常常看到，在一组条件之下，有多个随机事件．其中有些是比较简单的，也有比较复杂的．分析事件之间的关系，从而找到它们的概率以及概率之间的关系，这自然是必要的．而其基本点还是要搞清楚事件间的关系．

1. 事件的包含与相等

设有事件 A 及 B．如果 A 发生，那么 B 必发生，就称事件 B 包含事件 A，并记作

$$A\subset B \text{ 或 } B\supset A$$

例如投掷两枚匀称的分币，令 A 表示"正好一个正面朝上"，B 表示"至少一个正面朝上"，显然有 $A\subset B$.

如果事件 A 包含事件 B，同时事件 B 也包含事件 A，那么就称事件 A 与 B 相等(或称等价)，并记作

$$A=B$$

2. 事件的和与积

定义 事件"A 或 B"称为事件 A 与事件 B 的和，记作 $A+B$ 或 $A\cup B$；事件"A 且 B"称为事件 A 与 B 的积，记作 $A\cdot B$ 或 AB

或$A\cap B$. 这就是说，$A+B$发生，即“A或B”发生，它意味着A，B中至少有一个发生；$A\cdot B$发生，即“A且B”发生，它意味着A，B都发生.

例如，投掷两枚匀称的分币，A表示“正好一个正面朝上”的事件，B表示“正好两个正面朝上”的事件，C表示“至少一个正面朝上”的事件. 于是有

$$A+B=C,\quad AC=A$$

$$BC=B,\quad AB=V\text{(不可能事件)}$$

定义 事件“非A”称为A的**对立事件**，记作$\bar{A}$.

例如，投掷两枚分币，事件“至少一个正面朝上”是事件“两个都是正面朝下”的对立事件.

由该定义可知

$$\overline{(\bar{A})}=A$$

即A也是$\bar{A}$的对立事件. 我们看到：

在一次试验中，A和$\bar{A}$不会同时发生(即它们互相排斥)而且A，$\bar{A}$至少有一个发生. 就是说，A和$\bar{A}$满足：

$$\begin{aligned}A\cdot\bar{A}&=V\\ A+\bar{A}&=U\end{aligned}\tag{3.1}$$

定义 事件A同B的**差**表示A发生而B不发生的事件，记作$A-B$.

由上述定义可知

$$A-B=A\cdot\bar{B}\tag{3.2}$$

再举一个打靶的例子. 事件A代表命中图1.1 (a) 的小圆内，事件B代表命中图1.1 (b)的大圆内. 则$A+B$代表命中图1.1 (c)的阴影. $A\cdot B$代表命中图1.1 (d)的阴影，$A-B$代表命中图1.1 (e)的阴影.