黄午阳 编

概率统计初步

上海科学技术文献出版社

上海市职工高等学校试用教材

概率统计初步

黄午阳 编

上海科学技术文献出版社

上海市职工高等学校试用教材
概率统计初步
黄午阳 编
•
上海科学技术文献出版社出版
（上海市武康路2号）
新华书店上海发行所发行
上 海 商 务 印 刷 厂 印 刷
•
开本 787×1092 1/32 印张 5.25 字数 126,000
1984年3月第1版 1984年3月第1次印刷
印数：1—30,000

书号：13192·57 定价：0.66元

《科技新书目》63-241

前　言

《概率统计初步》适合于作工科大专或中专的教材，也可供自学者参考。全书包括随机事件与概率、随机变量、随机向量、参数估计、假设检验等内容，40学时左右可学完。考虑到学时数的限止，书中未涉及随机过程、方差分析、回归分析等内容。在内容的处理上，我们着重于讲清基本概念和基本方法，而对于一些较复杂的定理并未给出证明。

由于概率统计的解题方法（特别是基础概率部分）与其它数学学科相比，有些独特之处，因此初学者往往感到比较困难。为此，我们在基础概率部分适当增加了一些例题以供初学者参考。

在编写《概率统计初步》的过程中，主要参考了陈家鼎、刘婉如、汪仁官编的《概率统计讲义》，浙江大学数学系高等数学教研组编的《概率论与数理统计》以及梁文沛、黄午阳编的《初等概率论》等教材。在此，向他们表示感谢。由于编者水平有限，书中的缺点和差错一定不少，欢迎广大读者批评指正。

编　者

1983.3

目 录

第一章　随机事件与概率

概率论是数学的一个重要分支，它研究随机现象内部所蕴含的必然规律性。

在日常生活、生产实践和科学实验中，随机现象是大量存在的。比如：掷一枚重量均匀的硬币，静止后可能正面向上，也可能反面向上；医院中的产妇，可能生男孩，也可能生女孩；某天的天气可能是晴天，也可能是阴天或雨天；从一批产品中任取一件，取到的可能是正品，也可能是次品；某天黄浦江的水位；电话总机在一小时内收到的呼唤次数等等，这些现象都带有不确定性，也就是说有随机性。随机性仅仅是随机现象的一个方面，随机现象还有规律性的一面。概率论的任务，即研究它的规律性，及随机事件发生的可能性大小的量。

第一节　预备知识
——排列与组合

排列与组合是研究古典概率的重要工具。在此对这部分内容作一次系统的复习。

例1　从甲城到乙城，可以坐轮船，也可以坐汽车或火车。如一天中，轮船有三班，火车有四班，汽车有五班，从甲城到乙城共有多少种走法？

分析　因为坐轮船有3种走法，坐火车有4种走法，坐汽车有5种走法，每一种走法都可以从甲城到乙城。因此从甲城到乙城共有$3+4+5=12$种走法。

假设完成一件事有 m 类办法，在第一类办法中有 n_1 种方法，在第二类办法中有 n_2 种方法，……，在第 m 类办法中有 n_m 种方法，则完成这件事共有 $n_1+n_2+\cdots+n_m$ 种不同的方法。这就是加法原理。

例 2 从甲村到丙村必须经过乙村，已知甲村到乙村有二条路可走，乙村到丙村有三条路可走，问从甲村到丙村共有几种走法。

分析 因为从甲村到乙村共有 2 种走法，无论按那一种走法到乙村后，又有 3 种走法到丙村。所以从甲村到丙村共有 $3+3=2\times3=6$ 种走法。

例 3 有彩旗红、黄、绿、白 4 面，若连续举 3 面不同颜色的旗组成一个信号，一共可发出多少种信号？

分析 出第一面旗共有 4 种方法，而对于其中的每一种方法，出第二面旗有 3 种方法，对于这 3 种方法的每一种，出第三面旗又有 2 种方法，因此共有 $4\times3\times2=24$ 种不同的信号。

从分析上面的例子中可以看出，虽然例 2 和例 3 是两个不同的问题，但却具有共同的特点：

假设完成一件事有 m 个步骤，第一步有 n_1 种方法，第二步有 n_2 种方法，……，第 m 步有 n_m 种方法，且必须通过 m 个步骤才算完成这件事，则完成这件事共有 $n_1\times n_2\times\cdots\times n_m$ 种方法。这就是乘法原理。

例 4 由五个不同的自然数 1、2、3、4、5，可组成多少个没有重复数字的二位数。

解 根据乘法原理，组成二位数这件事有两个步骤，第一步是选十位数字，$n_1=5$，第二步是选个位数字，$n_2=4$，所以可组成 $5\times4=20$ 个没有重复数字的二位数。

例 3 讲的是由三面旗构成的信号，例 4 讲的是两个自然数

组成的二位数，如果撇开具体对象，把旗与自然数统称为元素，那末，例 3 可以这样叙述：

在四个不同的元素中任取三个，按一定的顺序排成一排，所有的排列种数有多少。

同样例 4 可叙述为：

在五个不同的元素中任取两个，按一定的顺序排成一排，所有的排列种数有多少。

定义 从 m 个元素中，每次取出 $n(n<m)$ 个按照一定的顺序排成一排，叫做从 m 个元素中每次取 n 个的选排列，所作出的不同排列的种数用 A_m^n 表示。

从例 3、例 4 可知 A_m^n 等于 n 个连续自然数的乘积，而最大的一个自然数为 m，即

$$A_m^n=m(m-1)\cdots(m-n+1) \tag{1.1}$$

例 5 由五个不同的自然数 1、2、3、4、5，可组成多少个没有重复数字的五位数。

解 这相当于从五个元素中，每次取出五个按照一定的顺序排成一排，由(1.1)可知，不同排列的种数(即没有重复的五位数)为：

$$A_5^5=5\times4\times3\times2\times1=120$$

把所有元素都取出，按照一定的顺序排成一排叫做全排列，这时用另一个记号 P_m(即 $P_m=A_m^m$)表示所有的排列种数，它是从 1 开始的 m 个连续自然数的乘积，称为 m 阶乘，记作 $m!$。以后我们所说的排列，一般是指选排列。

例 6 由五个不同的自然数 1、2、3、4、5，可组成多少个可有重复数字的二位数。

解 因为十位数字和个位数字允许重复，故选十位数字时有 5 种选法，选个位数字时仍有 5 种选法。由乘法原理可知，共

可组成 $5\times5=25$ 个有重复数字的二位数。

这种允许元素可重复出现的排列，叫做重复排列。

从 m 个不同的元素里，每次取出 n 个元素(元素可以重复出现，n 也不一定要小于或等于 m) 按照一定的顺序排成一排，那末第一，第二，……，第 n 位上选取元素的方法都有 m 种。所以，从 m 个不同的元素里，每次取出 n 个元素的重复排列的种数是:

$$\underbrace{m\times m\times\cdots\times m}_{n}=m^n \tag{1.2}$$

例 7 四个青年工人到照相馆去拍照，任选 3 人排成一排，拍摄一张照片，问有多少种拍法？还是这四个青年工人，从中任选 3 人组成一个技术革新小组，问有多少种选法？

解 把这四个青年工人看成 4 个不同的元素，并分别用 a, b, c, d 来表示，则第一个问题就是从 4 个元素中任取 3 个的排列问题，它的种数为 $A_4^3=4\times3\times2=24$。我们不妨把所有不同的排列列举在下面:

abc	abd	acd	bcd
acb	adb	adc	bdc
bac	bad	cad	cbd
bca	bda	cda	cdb
cab	dab	dac	dbc
cba	dba	dca	dcb

后面一个问题是从 4 个元素中任取 3 个元素，但不考虑它们之间的顺序排成一组，这叫做从 4 个元素中每次取 3 个的组合，所有不同组合的种数用 C_4^3 表示。不妨也把所有不同的组合列举出来:

abc　　abd　　acd　　bcd

这就是说 $C_4^3=4$。

这两个问题有什么异同之点呢？首先，它们都是从 4 个元素中任取 3 个元素，这是相同之点。但是前者与顺序有关，而后者与顺序无关。因此这两个问题既有区别 又有联系，不难看出，排列问题中的第一列的 6 种不同的排列，在组合问题中，因为不存在顺序关系，所以只组成一个组合；反之，如果把 abc 这个组合的 3 个元素，排成不同的顺序，它就有 $3!=6$ 种不同的排列，这也就是上面问题中的第一列的 6 种不同的排列。于是，不难得出组合种数 C_4^3 和排列种数 A_4^3 之间的关系如下：

$$A_4^3=3!C_4^3$$

或者

$$C_4^3=A_4^3/3!$$

一般说来，从 m 个元素中每次取 n 个元素的所有不同组合的种数

$$C_m^n=A_m^n/n! \tag{1.3}$$

关于组合有以下一些重要性质：

性质 1 $C_m^n=C_m^{m-n}$

证 因为

$$C_m^n=\frac{A_m^n}{n!}=\frac{m(m-1)\cdots(m-n+1)(m-n)!}{n!(m-n)!}$$

$$=\frac{m!}{n!(m-n)!}$$

$$C_m^{m-n}=\frac{A_m^{m-n}}{(m-n)!}=\frac{m(m-1)\cdots[m-(m-n)+1]n!}{(m-n)!n!}$$

$$=\frac{m!}{n!(m-n)!}$$

所以

$$C_m^n=C_m^{m-n}$$

性质 2 $C_m^n+C_m^{n-1}=C_{m+1}^n$

证 $$C_m^n+C_m^{n-1}=\frac{m!}{n!(m-n)!}+\frac{m!}{(n-1)![m-(n-1)]!}$$
$$=\frac{m!(m-n+1)+m!n}{n![(m+1)-n]!}$$
$$=\frac{m!(m+1)}{n![(m+1)-n]!}$$
$$=\frac{(m+1)!}{n![(m+1)-n]!}=C_{m+1}^n$$

第二节 随机现象

自然现象中，有一类现象是确定性的。比如：在一个大气压下，水加热到100°C就沸腾，冷却到0°C就结冰；知道了圆的半径，它的周长与面积就完全确定了；知道了某一物体的上抛初速度 v_0(如果不计空气阻力)，这个物体以后的运动状态就完全确定了。但是，也常常会遇到另一类现象，例如：掷一枚重量均匀的硬币，可能出现正面向上，也可能出现反面向上；医院中的产妇，可能生男孩，也可能生女孩等等。诸如此类不确定的现象，称为随机现象。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。

既然随机事件有偶然性，那么怎样去探讨和研究呢？例如：有50张相同的卡片，其中有一张写着“一等奖”，将50张卡片放在袋里，然后让一个人去抽这张写有“一等奖”的卡片，这时往往会听到这样的说法：“抽抽看，还有1/50的可能性。”这里，前半句话反映了事件的偶然性——随机性；而后半句话，却又反映了随机事件的必然性——规律性。如抽签的人第一次就抽到“一等奖”，是偶然的，即使抽到第三次、第四次，才抽到“一等奖”，也还是偶然的。可是，抽成千上万次，乃至数十万次，那末就会发现，

抽到的次数与抽的总次数之比总是接近1/50。这说明在同样条件下进行大量试验时，随机事件的规律性才呈现出来，这是在我们生活实践中所证实了的真理。又比如在一个医院里，某一天的产妇生男孩或女孩完全是偶然的，可是，在人口统计中，不论是那一个国家、何种民族，也不论是什么时期的统计资料，都发现有同样的规律：男孩的出生率是22/43。曾经发生这样一件事，法国某地男孩的出生率与22/43有较大的偏差，但经分析后发现是资料统计上存在问题，在正确地处理了原始资料后，男孩的出生率仍接近于22/43。

这些事例说明了随机事件的本质：在一次试验中，其结果是不确定的；即带有偶然性；但在大量的重复试验下，却又呈现出明显的规律性。这种规律性是通过对随机事件进行大量的试验后得到的。因此，称它为统计规律性。

在一组条件S之下，如果事件A一定发生，则称事件A为必然事件。必然事件通常用U表示。反之，在一组条件S之下，不可能发生的事件称为不可能事件，记作V。事件可能发生，也可能不发生，则为随机事件。这里条件组S很重要，初学时必须注意这一点，因为它影响着事件的性质。比如：水沸腾的条件组S，由两个条件构成：一是在一个大气压之下，二是要加热到100℃。如果忽略了一个大气压这个条件，即使加热到100℃也不一定会沸腾。这就说明，忽略了一个条件，必然事件就可能转化为随机事件。

第三节　概率的两种定义
——统计概型和古典概型

我们知道，随机事件在大量的重复试验下会呈现其规律性。

现在，再来看下面的一个例子。

为了确定某类种子的发芽率，从大批种子中抽出若干批做发芽试验，其结果如下：

种子粒数	25	10	70	130	310	700	1500	2000	3000
发芽粒数	24	9	60	116	282	639	1339	1806	2715
发芽率	0.96	0.9	0.875	0.892	0.910	0.913	0.893	0.903	0.905

从上面的数字可看出，发芽率在0.9附近摆动。这个例子说明，只要在同一组条件S之下作大量的重复试验，事件A发生的频率（事件A发生的次数与试验的总次数之比）就会呈现某种稳定性。一般说来，当试验的次数增加时，事件A发生的频率总是稳定于某一数p附近，而且偏离的幅度很微小。频率具有稳定性这一事实，说明刻划事件A发生的可能性大小的量——概率，是客观存在的。

定义 在固定的一组条件S之下，重复作n次试验，当试验的次数n很大时，如果事件A发生的频率μ/n稳定地在某一数值p的附近摆动，而且，一般说来，随着试验次数的增加，这种摆动的幅度是很微小的，则称A为随机事件，p为随机事件A在条件组S之下发生的概率，记作：

$$P(A)=p$$

这就是概率的统计定义，简称统计概型。根据概率的统计定义，容易得到：$P(U)=1$, $P(V)=0$。

由统计概型求某事件A的概率的方法乃是试验的方法。在实际工作中，精确的p值常常是无法求得的，因此，当试验次数n适当大时，我们就把频率μ/n作为p的近似值。

有一类随机现象最简单却又是常见的。例如：掷一枚重量均匀的硬币，当它静止后总是正面朝上或反面朝上。由于硬币是对称的几何体，所以出现正面与反面的可能性是相等的，即

都等于1/2。此外，每掷一次后不是正面朝上就是反面朝上且二者不可能同时发生。又如，掷一粒均匀的骰子，哪一面朝上有六种可能，且出现这六种点中的任何一点的可能性是相等的，即都等于1/6。此外，每掷一次，六种点中至少有一种点出现，且至多也只有一种点出现。

上述随机现象是概率论早期所研究的对象，所以称为古典概型。对这类随机现象进行归纳，可得出如下的一般规律：

如某种试验只有有限种不同的结果，即只可能出现有限个事件 $A_1, A_2, \cdots, A_n$，且它们具有下列三条性质：

(1) $A_1, A_2, \cdots, A_n$ 出现的可能性相等(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(完备性)；

(3) 在任一次试验中，$A_1, A_2, \cdots, A_n$ 至多有一个发生(互不相容性)

则事件组 $A_1, A_2, \cdots, A_n$ 称为一个等可能完备事件组或等概基本事件组，其中任一事件 $A_i(i=1, 2, \cdots, n)$ 称为基本事件。

等可能性、完备性和互不相容性是古典概型中基本事件的主要特征。下面给出古典概型的计算公式。

若 $A_1, A_2, \cdots, A_n$ 是一个等概基本事件组（即基本事件的总数为 n），而事件 A 由其中的 m 个基本事件所组成（或者说只有 m 个基本事件有利于 A），则事件 A 的概率为：

$$P(A)=\frac{m}{n} \tag{3.1}$$

因此，在解古典概型问题时，只要求出基本事件的总数和有利事件数即可。

例1 袋中有5个白球和3个黑球，从中任意取出2个球，问取出的2个都是白球的概率是多少？

$$\frac{C_5^2}{C_8^2}$$

解 袋中共有 8 个球，从 8 个球中任取 2 个球的方法有 $C_8^2 = \frac{8\times7}{2!} = 28$ 种，即基本事件的总数 n 为 28。而从 5 个白球中任取 2 个的方法有 $C_5^2 = \frac{5\times4}{2!} = 10$ 种，即有利事件数 m 为 10。故取出的 2 个都是白球的概率

$$p = \frac{10}{28} = \frac{5}{14} = 0.357$$

$C_5^3 C_3^1$

例 2 在上例中，如从袋中任取 4 球，问恰好抽到 3 个白球的概率是多少？

解 基本事件总数 n 为 $C_8^4 = \frac{8\times7\times6\times5}{4!} = 70$。

因为取出的 4 球中，恰有 3 个白球时另一球必为黑球。3 个白球必是从 5 个白球中取出的，共有 C_5^3 种；而 1 个黑球必是从 3 个黑球中取出的，共有 C_3^1 种。因此，有利事件数 m 为 $C_5^3 \cdot C_3^1 = 30$。故恰好抽到 3 个白球的概率

$$p = \frac{C_5^3 C_3^1}{C_8^4} = \frac{3}{7}$$

值得指出的是，解古典概型问题时要特别注意事件的等可能性，否则往往容易出错。例如，两袋中分别装有写着 0, 1, 2, 3, 4, 5 六个数字的六张卡片，从每袋中各取一张，求所得卡片上两数之和等于 6 的概率。对于这个问题，某学员是这样解的：两数之和共有 0, 1, 2, 3, …, 10 共十一种不同结果，即基本事件的总数为 11，而有利事件数为1，故所求的概率为 1/11。请大家思考一下这种解法错在哪里，正确的答案应是多少。

第四节 事件的关系与运算

世界上的事物不是孤立的，而是相互联系、相互制约的。因

此，当研究事件的概率时，也应该从它们的联系中去思考问题、分析问题。弄清所研究的事件的关系及运算对于推导公式、计算概率都会带来很大的方便。

先看一个例子。在检验一批圆柱形的产品时，要求产品的长度和直径都合规格才算合格，这时就得考虑"产品合格"、"产品不合格"、"长度合格"、"长度不合格"、"长度不合格而直径合格"、"直径不合格而长度合格"等事件。

定义 设有事件 A 及 B，若 A 发生必然导致 B 发生，称事件 B 包含事件 A，记作 $B\supset A$ 或 $A\subset B$。这种关系如图 1.1 所示。

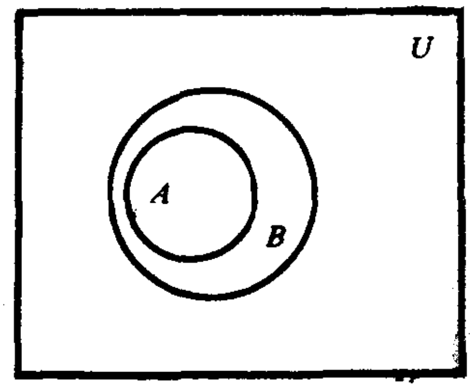

图 1.1

例如：检验圆柱形产品时，"长度不合格"(A) 必然导致"产品不合格"(B)，于是 $B\supset A$。

定义 若事件 B 包含事件 A，同时事件 A 又包含事件 B，称事件 A 与事件 B 相等（等价），记作 $A=B$。

定义 若事件 A 与事件 B 中至少有一个发生的事件，称为事件 A 与 B 的和（并），记作 $A+B$ 或 $A\cup B$。A 与 B 的和如图 1.2 所示。

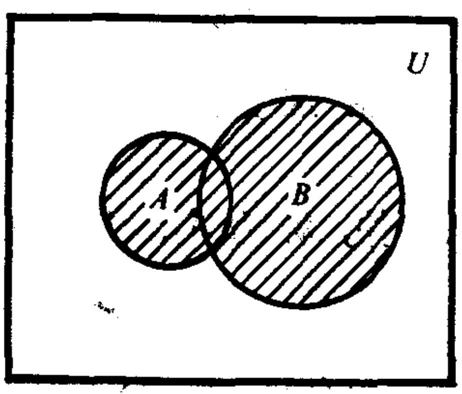

图 1.2

例如：在检验圆柱形产品时，"产品不合格"这个事件便是"直径不合格"(A) 与"长度不合格"(B)之和。

事件 $A_1, A_2, \cdots, A_n$ 中至少有一个发生的事件，称为 $A_1, A_2, \cdots, A_n$ 的和（并），记作