中等专业学校统计专业试用教材

数理统计

周鼎权 主编

中国统计出版社

中等专业学校统计专业试用教材

数 理 统 计

周鼎权 主编

中国统计出版社

中等专业学校统计专业试用教材
数　理　统　计
周鼎权　主编

中国统计出版社出版发行
北京北方印刷厂印刷
北京印刷一厂排版
787×1092毫米　32开　9.375 印张　20万字
1985年6月第1版　1985年6月北京第1次印刷
印数：1—31,000
统一书号：4006·070　定价：2.00元

出版说明

为了适应中等专业学校统计专业教学的需要，国家统计局人事教育司组织编写了一套专业课试用教材，这本《数理统计》是其中之一。本书由四川统计学校讲师周鼎权、教师刘全、陈启富共同执笔，周鼎权主编。由国家统计局人事教育司教育处审定。

在本书的编写过程中，西安统计学院教师洪含淳、张恩祥，河南省计划统计学校教师李志远，北京市计划统计学校教师王曲均等同志参加了初稿的讨论，提出了许多宝贵的意见，特此致谢。

由于编写时间仓促，加上水平有限，书中的缺点和错误实属难免，恳请读者予以指正。

国家统计局人事教育司

一九八五年一月

目　录

第一章 基础概率

第一节 随机事件及其概率

一、确定性现象与随机现象

在自然界和社会生活中，有许多现象我们完全可以预先断定它们在一定条件下是否会出现。例如："同性的电互相排斥"，"在标准大气压下，纯净的水加热到 100°C 时必定会沸腾"等等，是一定会出现的。而上述现象的反面，即"同性的电互相吸引"，"在标准大气压下，纯净的水受热到 100°C 时不沸腾"等等，是必然不会出现的。这类现象有着一个共同的特点，就是事先对其发生与否可以做出肯定性的回答。这种在一定条件下必然发生或必然不发生的现象，称为确定性现象。

然而自然界和社会生活中，还广泛地存在着与上述确定性现象有着本质区别的另一类现象。例如随机地投掷一枚均匀的硬币，其结果可能是正面朝上，也可能是背面朝上；从一批既有正品也有次品的产品中任抽 4 件，这 4 件中可能有次品，也可能没有次品；等等。它们有着与确定性现象不同的特点，即在一定条件下，有多种可能的结果发生，事前人们不能预言将出现哪种结果，即呈现出不确定性。这种在一定条件下，可能发生，也可能不发生的现象称为随机现象。

人们经过长期实践和深入研究以后发现，尽管就个别的试验或观察而言，究竟会出现什么结果不能预先断定，即随机现

象的发生与否表现出不确定性的一面；但是，当我们进行大量重复试验或观察时，就会发现随机现象的发生与否呈现出某种固有的规律性，如多次重复抛一枚硬币得到正面向上大致有半数等。这种规律称为统计规律。

概率论和数理统计就是研究和揭示随机现象统计规律性的一门数学学科。

二、随机事件与概率

研究随机现象离不开随机试验。一般地，设E为一试验，如果事先不能准确地预言它的结果，而且在相同条件下可以重复进行，就称为随机试验。

为方便起见，以下用字母E表示随机试验。

在随机试验E中，每一个可能出现的最简单的结果ω，称为随机试验E的基本事件，全体基本事件的集合称为基本事件空间。记为$\Omega=\{\omega\}$。在具体问题中，十分重要的是，认清基本事件集是由什么构成的。下面举一些例子。

例 1-1 设E_1—将一枚均匀硬币投掷一次，观察所出现的面。ω_1—正面，ω_2—反面，于是Ω由两个基本事件构成，即$\Omega=\{\omega_1, \omega_2\}$。

例 1-2 设E_2—将一枚均匀硬币连续投掷二次，观察所出现的面。ω_1—正正，ω_2—正反，ω_3—反正，ω_4—反反，于是有$\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4\}$。

例 1-3 设E_3—从颜色分别为红、白、黑的三个球中不放回地抽取两个，观察各种颜色出现的先后秩序。ω_1—红白，ω_2—红黑，ω_3—白红，ω_4—白黑，ω_5—黑红，ω_6—黑白。于是有$\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$。

有了基本事件空间的概念，就可以定义一般的随机事件。我们定义随机事件为基本事件空间Ω的某个子集。我们把随机事件简称事件，常用A，B，C，…表示。

例 1-2 中，随机试验E_2的基本事件空间为$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$，事件$A_1$—“连续投掷二次中至少出现一次反面”是由3个基本事件组成，即$A_1=\{\omega_2,\omega_3,\omega_4\}$。事件$A_2$—“连续投掷二次均为正面”是由1个基本事件组成，即$A_2=\{\omega_1\}$。

例 1-3 中，随机试验E_3的基本事件空间是$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$。

子集$B_1=\{\omega_1,\omega_2\}$表示“第一次抽中红球”这一事件；

子集$B_2=\{\omega_1,\omega_2,\omega_3,\omega_5\}$表示“抽到红球”这一事件；

子集$B_3=\{\omega_6\}$表示“第一次抽中黑球，第二次抽中白球”这一事件。

基本事件是事件的一种，并且是不可再分割的事件，一般事件是由若干个基本事件共同组成。每一事件对应基本事件空间Ω的一个子集，当且仅当事件A所含的一个基本事件发生，事件A才发生。对于例 1-2 中的事件A_1，若试验的结果是ω_2，ω_3，ω_4之一，则事件A_1发生；对于例 1-3 中的事件B_2，若事件B_2发生，则试验的结果必是ω_1、ω_2、ω_3、ω_5之一。

为讨论方便，我们把基本事件空间Ω也作为一个事件，因为在每次试验中必然有Ω中的某个基本事件发生，即事件Ω在每次试验中必然发生。所以称Ω为必然事件。类似地，我们把不包含任何基本事件的空集也作为一个事件，由于它在每次试验中都不会发生，故称为不可能事件，记为Φ。

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，但当我们重复进行这一试验时，常常会察觉事件发生的可能性彼此是不相同的。这种"事件出现的可能性大小"是随机事件本身固有的属性。例如掷一枚骰子，"出现偶数点"要比"出现点2"的可能性大。为了研究随机事件的这种属性，我们先引入频率的概念。

设随机事件 A 在 n 次试验中出现 A_n 次，则比值

$$W_n\{A\}=\frac{A_n}{n}$$

称为事件 A 在这 n 次试验中出现的频率。

显然，$W_n\{\Omega\}=1$，$W_n\{\Phi\}=0$。对任何事件 A 都有

$$0\leqslant W_n\{A\}\leqslant 1$$

现在我们看看大量试验之后，随机现象有什么样的统计规律性。

例 1-4　一颗棉花种籽播种后，由于种种原因可能发芽也可能不发芽，所以"种籽发芽"是随机事件。当我们对数百粒种籽的发芽情况进行观察——即作了数百次试验，表面上看，"种籽发芽"事件仍是混杂的，毫无规律的，然而深入下去对这些试验结果进行分析，可以发现客观上存在着一定的规律。下面是棉花种籽在温室内的发芽试验的结果。

表 1-1

试验种籽数 n	50	100	150	200	250	300	350	400	450	500
发芽种籽数 A_n	27	55	85	117	152	184	201	234	269	298
发芽的频率 $\frac{A_n}{n}$	0.54	0.55	0.57	0.59	0.61	0.61	0.57	0.59	0.60	0.60

从表中明显地可以看出：随着 n 的增大，发芽频率 $\frac{A_n}{、}$ 呈现稳定性，即频率在一个定数(比如0.60)上下摆动；随着 n 的继续增大，出现与 0.60 有较大偏差的情况越来越少。

例 1—5 投掷一枚均匀的硬币，就每次投掷来说“出现正面”还是“出现反面”，都是偶然的。即“出现正面”或“出现反面”是随机事件。但是当我们在相同条件下重复进行多次试验时，也可以发现存在着一定的规律性。表 1-2 所列的是历史上一些比较著名的试验结果：

表 1-2

实验者	投掷次数	出现正面向上的次数	频率(百分比) $\frac{A_n}{n}$
De Morgan	2,048	1,061	0.518
Buffon	4,040	2,048	0.507
Pearson	12,000	6,019	0.5016
Pearson	24,000	12,012	0.5005

投掷硬币是毫无实际意义的，然而例子本身却说明了在“偶然”性中起着支配作用的“必然”规律是客观存在的。“出现正面”这一随机事件在 n 次重复的试验(在同样条件下投掷硬币)中出现的频率是在 0.5 这个定数附近摆动的。随着 n 增加，这种稳定在 0.5 这个数值附近的趋势越来越明显。这是一个非常重要的事实。

类似的例子，还可以列举不少。归纳起来，上述实例指明了随机现象遵循着一个客观规律，即频率的稳定性。

频率的稳定性：当试验次数 n 充分大时，事件 A 发生的频率 $\frac{A_n}{n}$ 在一个定数上下摆动，并且随着 n 的增加，这个频率会越来越接近这个定数。

频率的稳定性指明了随机现象中存在的规律，它说明，随机事件就每一次试验而言是可能发生也可能不发生的，然而在大量的重复试验中，事件发生的可能性大小是具有客观规律的。

显然，若随机事件的频率稳定在一个较大数值附近，表明该随机事件发生的可能性大；若随机事件的频率稳定在一个较小数值附近，则表明该事件发生的可能性小。可见，频率所接近的那个固定的数值是相应事件发生的可能性大小的一个客观存在的量度。我们把具有这种性质的数字称为相应事件的概率。

概率的统计定义：在一组不变的条件 S 下，重复作 n 次试验，A_n 是 n 次试验中事件 A 发生的次数，当 n 充分大时，如果频率 $\frac{A_n}{n}$ 稳定地在某一常数 p 的上下摆动，而且一般说来随着 n 的增大，这种摆动的幅度会越来越小，则称 p 为随机事件 A 在该条件组下发生的概率。记作

$$P(A)=p$$

事件的频率与事件的概率是两个不同的概念，二者有本质的区别。事件的频率是与进行的试验有关的一个相对数 $\frac{A_n}{n}$，是随着试验不同而不同的。而事件的概率反映的是随机现象的某种本质属性，是与试验次数无关而客观存在的一个确定的数

值。频率是概率的表现形式，概率是现象的本质属性，决定着频率的变动趋势。

概率的统计定义给出了实际求事件A的概率p的方法。但有时这个定数p不易求得，人们就采用在大量试验之后得到的事件频率作为它的概率的近似值，并称它为经验概率。由频率的稳定性保证了这种近似代替的合理性。我们实际上正是这样看待问题的。比如长度的概念，并不因为每次实测数值都是近似的而建立不起来，也不会因为温度计读数都是近似的而怀疑“温度”的客观存在。

由于频率$\frac{A_n}{n}$总介于0到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Ω和不可能事件Φ，显然有

$$P(\Omega)=1, \quad P(\Phi)=0.$$

第二节　古 典 概 型

一、古典概型的概念

在各种随机试验中，有一类最简单但却常见的随机试验，这类随机试验有以下特征：

(1) 只有有限个不同的基本事件ω_1，ω_2，…，ω_n，即基本事件空间$\Omega=\{\omega_1, \omega_2, \cdots, \omega_n\}$；

(2) 所有的基本事件都是等可能的。

具有以上两个特征的随机试验称为古典型的随机试验。

概率的古典定义：对古典型的随机试验E，若基本事件空

间 $\Omega=\{\omega_1, \omega_2, \cdots, \omega_n\}$，事件 A 由 $m(m\leqslant n)$ 个不同的基本事件组成，我们定义事件 A 出现的概率为：

$$P(A)=\frac{A\text{中所含的基本事件数}}{\text{基本事件总数}}=\frac{m}{n}$$

这样，在古典型的随机试验下，确定一个事件 A 的概率的问题，就可化为计算总的基本事件的个数以及组成事件 A 的基本事件的个数之比的问题。

对于这类随机现象的概率数学模型，习惯上也称为古典概型。

显然，在古典概型中，事件的概率仍具有以下性质：

(i) 对任何事件 A，$0\leqslant P(A)\leqslant 1$；

(ii) $P(\Omega)=1$，$P(\Phi)=0$。

二、古典概型的计算

根据古典概率定义，我们可以通过分析直接求得在古典型的随机试验 E 下的事件的概率，而无需做大量的重复试验。需要指出的是，在分析一个随机现象时，必须首先弄清在一个试验中（即条件组实现之下）所有可能的基本事件是什么，以及所讨论的事件是由哪些基本事件组成的。这在古典概型中确定事件的概率是非常重要的。

例 1-6　将一枚均匀硬币任意投掷一次，求出现正面的概率。

解：设 $A=$“投掷一次恰出现正面”

显然，基本事件空间 $\Omega=\{$正、反$\}$。

$A=\{$正$\}$

故　$P(A)=\frac{1}{2}$

由于每次试验（投币）中出现正面或出现反面都是等可能的，且每次试验能且只能出现二者之一，可以设想，当我们作大量重复投币试验后，事件“出现正面”的频率是会稳定在 1/2 左右的。故由概率的统计性定义知事件“出现正面”的概率为 1/2 是合理的。

例 1-7　袋中装有 3 个红球、2 个白球和 1 个黑球，现从袋中任取 1 球，求以下事件的概率：

A =“取得红球”；

B =“红球不被取出”。

解：设想将 6 个球分别编号，其中红球为(1)、(2)、(3)；白球为(4)、(5)；黑球为(6)。

显然，$\Omega=\{(1),(2),(3),(4),(5),(6)\}$；

$A=\{(1),(2),(3)\}$；

$B=\{(4),(5),(6)\}$；

故

$$P(A)=\frac{3}{6}=\frac{1}{2}$$

$$P(B)=\frac{3}{6}=\frac{1}{2}$$

例 1-8　盒中装有 2 个红球，1 个白球，从盒中随机抽取 2 球，就下列几种情形

(1) 抽取无先后之分(可以考虑一次抽出两球)

(2) 抽取有先后，每次抽一球，抽后不放回

(3) 抽取有先后，每次抽一球，抽后放回

分别计算事件 A =“取得一个红球一个白球”的概率。

解：将球编号：红球－1、红球－2、白球－3。

在情形(1)下：

基本事件空间 $\Omega=\{(1,\ 2)\ (1,\ 3)\ (2,\ 3)\}$

$$A=\{(1,\ 3)\ (2,\ 3)\}$$

故 $P(A)=\dfrac{2}{3}$。

在情形(2)下：

基本事件空间 $\Omega=\{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}$；

$$A=\{(1,3),(2,3),(3,1),(3,2)\}。$$

故 $P(A)=\dfrac{4}{6}=\dfrac{2}{3}$

在情形(3)下：

基本事件空间 $\Omega=\{(1,1),(2,2),(3,3),(1,2),(1,3)$ $(2,1),(2,3),(3,1),(3,2)\}$；

$$A=\{(1,3),(2,3),(3,1),(3,2)\}。$$

故 $P(A)=\dfrac{4}{9}$。

可见，由古典概型定义求事件的概率时，必须弄清楚问题中“试验”是什么，此“试验”下的基本事件又是什么。“试验”不同(即条件组不同)，基本事件也就不同，从而所求事件的概率也不相同。

例 1-9　从装有 7 个红球、3 个白球的盒中任意取 5 球，试求抽得的 5 球中恰有 3 个红球 2 个白球的概率。

解：设 $A=$“抽得的 5 球中恰有 3 个红球 2 个白球”

据题意可知，此处的“试验”是从 10 个球中任取 5 个，该

“试验”下的基本事件是 10 个球中任意 5 个球的组合，故基本事件空间由 C_{10}^5 个基本事件组成。

7 个红球中的任意 3 个和 3 个白球中的任意 2 个组成 A 的一个基本事件，故 A 由 $C_7^3C_3^2$ 个基本事件组成，则

$$P(A)=\frac{C_7^3C_3^2}{C_{10}^5}=0.42$$

第三节　事件的关系及其运算

概率论的重要研究课题之一是从简单事件的概率推算出复杂事件的概率。在实际问题中，往往要求我们同时考察几个在同样试验条件下的事件及它们之间的联系。详细地分析事件之间的关系，这不仅有助于我们更深刻地认识事件的本质，而且可以大大简化一些复杂事件的概率计算。

一、事件的关系及运算

设试验 E 的基本事件空间为 Ω，A、B、C、A_k $(k=1,2,\cdots)$ 为试验 E 所包含的事件。

我们已经知道，事件 A 是基本事件空间 Ω 中的一个子集，这子集仍然记为 A。那么在讨论事件之间的关系及运算时，可以从集合论的观点来理解这些问题。事件间的关系与集合间相应的关系是一致的。

（一）包含关系

若 A 发生则必有 B 发生，则称 B 包含 A。记为 $A\subset B$（或 $B\supset A$）。这种关系如图 1-1 所示。

（二）等价关系