概率论与数理统计

李　波　　吴渝春　编

重庆大学出版社

概率论与数理统计

李　波　　吴渝春　编

重庆大学出版社

内容简介

本书是针对农科院校学生的特点编写的，主要介绍了概率论与数理统计的基本概念、基本定理、基本方法. 本书注重概念的引入、定理的实际应用价值，并在例题和习题的选择上尽量选用农科问题中的一些实例. 在第十章中给出了概率论与数理统计中常见问题的计算机程序. 本书主要包括事件与概率、一维随机变量的分布函数、多维随机变量、随机变量的数字特征、大数定律与中心极限定理、点估计、假设检验、方差分析与回归分析、统计程序设计等内容.

本书可作为农科院校学生的教材，也可供其他专业学生参考.

概率论与数理统计

李　波　吴渝春　编

责任编辑　刘茂林

*

重庆大学出版社出版发行

新　华　书　店　经　销

重庆通信学院印刷厂印刷

*

开本：787×1092 1/32　印张：6.25　字数：140千

1995年3月第1版　1995年3月第1次印刷

印数：1-4500

ISBN 7-5624-1019-4/O·118　定价：6.50元

（川）新登字020号

目　录

43186

第一章　事件与概率

第一节　随机现象与统计规律性

概率论与数理统计是研究随机现象的数量规律的数学分支．什么是随机现象呢？先看两个例子．

例 1　一个袋子中有 10 个完全相同的白球，搅匀后从中任意摸取一球．

例 2　一个袋子中有 10 个相同的球，但 5 个是白色的，另外 5 个是黑色的，搅匀后从中任意摸取一球．

对于例 1，在球没有取出之前，就能确定取出的必定是白球．这种问题，根据开始的条件，就可以确定结果．而对于例 2 来讲，在球没有取出以前，从开始的条件就知道，不能确定结果即取出的球是白的还是黑的．也就是说一次摸取的结果，出现白的还是黑的，在摸取之前是无法确定的．对于这类问题，似乎没有什么规律可言．但是，由实践得知，如果从袋子中反复多次取球（每次取出一球，记录球的颜色后仍把球放回袋子中并且搅匀），那么总可以观察到这样的事实：当试验次数 n 相当大时，出现白球的次数 $n_{白}$ 和出现黑球的次数 $n_{黑}$ 是很接近的，比值 $\frac{n_{白}}{n}$（或 $\frac{n_{黑}}{n}$）会逐渐稳定于 1/2．出现这个事实是完全可以理解的，因为袋子中白球数等于黑球数，从中任意取一个球，取得白球或黑球的“机会”应该是平等的．

可见上面两个例子代表了两种类型的问题．例 1 所代表

的类型，在问题开始就能判定一个确定的结果，这种类型的问题所对应的现象称为确定性现象．确定性现象非常广泛，例如：

“早晨，太阳必然从东方升起.”

“苹果，不抓住必然往下掉.”

“边长为 a、b 的矩形，其面积必为 $a \cdot b$.”

例 2 所代表的类型，它有多于一种可能的结果，但是在一次摸取之前不能肯定会出现哪一个结果，就一次摸取而言，看不出什么规律性．但是，“大数次”地重复抽取，抽取的结果就会遵循某些规律，这种规律称之为”统计规律”，这类问题称为随机试验，试验所代表的现象称为随机现象．现实中随机现象是极为普遍的，例如：

“某地区的年降雨量.”

“检查一件产品，是合格品还是不合格品.”

“某电话交换站在单位时间内收到用户呼唤的次数.”

概率论与数理统计就是研究随机现象的统计规律的数学学科．由于随机现象的普遍性，使得它的应用极其广泛．近年来，由于计算机的普及，使得它为科学技术、工农业生产的现代化作出了重要的贡献．

第二节　随机事件与样本空间

在第一节中已介绍了随机试验，现进一步明确它的含意．如果一个试验满足下述条件：

(1) 试验可以在相同的情形下重复进行；

(2) 试验的所有可能结果是明确可知道的，并且不止一个；

(3) 每次试验总是恰好出现这些可能结果中的一个，但在一次试验之前却不能肯定这次试验会出现哪一个结果.

称这样的试验是一个随机试验，简称为试验.

随机试验的每一个可能结果，称为基本事件. 因为随机试验的所有结果是明确的，从而所有的基本事件也是明确的，它们的全体，称作样本空间，通常用字母 Ω 表示. Ω 中的点，即基本事件，也称作样本点，通常用 ω 表示.

例 1 在第一节中的例 2 中，令

$$\omega_1 = \{取得白球\}, \omega_2 = \{取得黑球\}$$

则
$$\Omega = \{\omega_1, \omega_2\}.$$

例 2 一个袋子中有 10 个完全相同的球，分别标以号码 1，2，…，10，从中任取一球，令

$$i = \{取得球的标号为 i\}$$

则
$$\Omega = \{1, 2, \cdots, 10\}.$$

例 3 测量某地水温，令

$$t = \{测得的水温为 t℃\}$$

则
$$\Omega = \{0, 100\}.$$

在随机试验中，有时关心的是带有某些特征的基本事件是否发生. 如在例 2 中，可以研究

$$A = \{球的标号 = 6\},$$
$$B = \{球的标号为偶数\},$$
$$C = \{球的标号 \leqslant 5\}.$$

这些结果是否发生？其中 A 为一个基本事件，而 B 和 C 则由多个基本事件组成，相对于基本事件，就称它们为复杂事件. 无论是基本事件还是复杂事件，它们在试验中发生与否，都带有随机性，所以都称作随机事件或简称为事件. 习惯上人们常用大写字母 A、B、C 等表示事件. 在试验中，如果出现 A 中所

包含的某一个基本事件 ω，则称作 A 发生，并记为 $\omega\in A$.

样本空间 Ω 包含了全体基本事件，而随机事件不过是有某些特征的基本事件所组成，所以从集合论的的观点来看，一个随机事件不过是样本空间 Ω 的一个子集而已.

又因为 Ω 是所有基本事件所组成，因而在任一次试验中，必然要出现 Ω 中某一基本事件 ω，即 $\omega\in\Omega$，也就是说在试验中，Ω 必然会发生，所以今后又用 Ω 来代表一个必然事件. 相应地，空集 ϕ 可以看作是 Ω 的子集，在任一次试验中，不可能有 $\omega\in\phi$，即 ϕ 永远不可能发生，所以 ϕ 是不可能事件. 必然事件与不可能事件的发生与否，已经失去了"不确定性"，因而本质上它们不是随机事件.

一个样本空间 Ω 中，可以有许多的随机事件. 概率论的任务之一，是研究随机事件的规律，通过对比较简单事件规律的研究去掌握更复杂事件的规律. 为此，需要研究事件之间的关系和事件之间的一些运算.

如果没有特别的申明，在以下的叙述中总认为样本空间 Ω 已经给定，并且还给定了 Ω 中的一些事件，如 A、B、$A_i(i=1,2,\cdots)$等.

1. 如果事件 A 发生必然导致事件 B 发生，则称 B 包含了 A. 记作 $A\subset B$ 或 $B\supset A$. 比如前面提到过的 $A=\{$球的标号$=6\}$，这一事件发生就导致事件 $B=\{$球的标号是偶数$\}$的发生. 因为摸到标号为 6 的球意味着标号为偶数的球出现了.

可以给上述的含意以一个直观的几何解释，设样本空间 Ω 是一个正方形，A 和 B 是两个事件，也就是说是 Ω 的某两个子集. "A 发生必然导致 B 发生"意味着"属于 A 的 ω 必然属于 B"，即 A 中的点全在 B 中. 其几何图形如图 1.1 所示.

因为不可能事件 ϕ 不含有任何 ω，所以对任一事件 A，约

定 $\phi \subset A$.

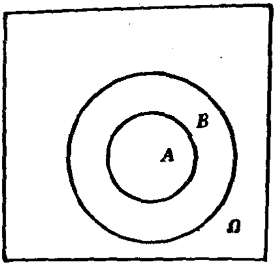

图 1.1

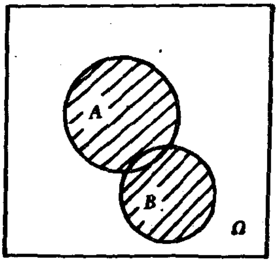

图 1.2

2. 如果 $A \subset B$ 和 $B \subset A$ 同时成立，则称事件 A 与事件 B 相等，记作 $A=B$. 易知，相等的两个事件 A、B，总是同时发生或同时不发生. 直截了当地说，所谓 $A=B$，就是 A，B 中包含有相同的样本点.

3. "事件 A 与 B 至少有一个发生"，这样的一个事件称作事件 A 与 B 的并，并记为 $A \cup B$. 它的几何表示见图 1.2. 图中的阴影部分为事件"$A \cup B$".

4. "事件 A 与 B 同时发生"，这样的事件称作 A 与 B 的交，记为 $A \cap B$（或 AB）. 它对应图 1.3 中的阴影部分.

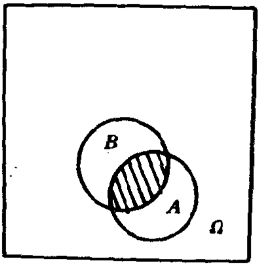

图 1.3

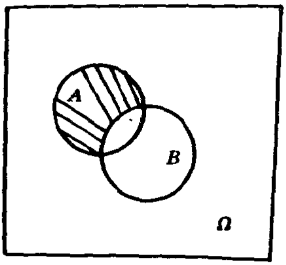

图 1.4

5. “事件 A 发生而 B 不发生”，这样的事件称为事件 A 与 B 的差，记作 $A-B$，它表示为图 1.4 中的阴影部分.

6. 若事件 A 与 B 不能同时发生，也就是说 AB 是一个不可能事件，即 $AB=\phi$，则称 A 与 B 互不相容，图 1.5 表示了这一情形.

7. 若 A 是一个事件，令 $\overline{A}=\Omega-A$，称 $\overline{A}$ 是 A 的对立事件或逆事件. 易知，在一次试验中，若 A 发生，则 $\overline{A}$ 必不发生(反之亦然)即 A 与 $\overline{A}$ 二者只能发生其中之一，并且也必然发生其中之一. 因而有

$$A\overline{A}=\phi, A\cup\overline{A}=\Omega, \overline{\overline{A}}=A$$

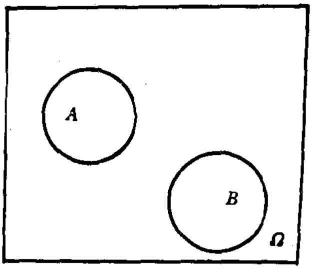

图 1.5

8. 若有 n 个事件：$A_1, A_2, \cdots, A_n$，则称“$A_1, A_2, \cdots A_n$ 中至少发生其中的一个”这样的事件称作 $A_1, A_2, \cdots, A_n$ 的并，并记作 $A_1\cup A_2\cup\cdots\cup A_n$ 或 $\bigcup\limits_{i=1}^{n}A_i$；若“$A_1, A_2, \cdots, A_n$ 同时发生”，这一事件称作 $A_1, A_2, \cdots, A_n$ 的交，并记作 $A_1A_2\cdots A_n$ 或 $\bigcap\limits_{i=1}^{n}A_i$.

如果读者有了高中数学中的集合论知识，一定会发现事件间的关系及运算与集合论中集合间的关系及运算之间是完全可以互相类比的. 下面给出这种类比的对应关系：

概率论	集合论
样本空间	$\Omega=\{\omega\}$
事件	子集
事件 A 发生	$\omega\in A$
事件 A 不发生	$\omega\notin A$

必然事件	Ω
不可能事件	ϕ
事件 A 发生导致事件 B 发生	$A\subset B$
“事件 A 与 B 至少有一个发生”	$A\cup B$
“事件 A 与 B 同时发生”	$A\cap B$(或 AB)
“事件 A 发生而 B 不发生”	$A-B$
事件 A 与 B 互不相容	$A\cap B=\phi$

在许多场合,用集合论的表达方式显得简练,也更容易理解.但对于初学概率论的人来说,重要的是要学会用概率论的语言来解释集合间的关系及运算,并能运用它们.

例 4 设 A、B、C 是 Ω 中的随机事件,则事件“A 与 B 发生,C 不发生”可以表示成 $AB\overline{C}$;“A、B、C 中至少有二个发生”可以表示成 $AB\cup AC\cup BC$;“A、B、C 中恰好发生二个”可以表示成 $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$;“$A$、$B$、$C$ 中有不多于一个事件发生”可以表示成:$\overline{A}\,\overline{B}\,\overline{C}\cup\overline{A}\,\overline{B}C\cup\overline{A}B\overline{C}\cup A\overline{B}\,\overline{C}$.

事件的运算满足下述规则:

(1)交换律:$A\cup B=B\cup A$,$AB=BA$;

(2)结合律:$(A\cup B)\cup C=A\cup(B\cup C)$;

$$(AB)C=A(BC);$$

(3)分配律:$(A\cup B)\cap C=AC\cup BC$,

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

(4)德摩根定理:$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A}_i$;

$$\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A}_i.$$

从前面已经知道事件是 Ω 的某些子集,如果把“是事件”的这些子集归在一起,则得到一个集合,记作 $\mathscr{F}$,称作事件域,即

$$\mathscr{F}=\{A:A\subset\Omega, A\text{是事件}\}$$

第三节　概率和频率

回忆第一节中的例 2，已知道它是一个随机试验，并且样本空间 $\Omega=\{\omega_1,\omega_2\}$，其中 ω_1 {取得白球}，$\omega_2=$ {取得黑球}是基本事件．在一次试验中，虽然不能肯定是 ω_1 还是 ω_2 发生，但是可以问在一次试验中发生 ω_1（或 ω_2）的可能性有大？由对称性，很自然地可以断定在一次试验中，出现 ω_1（或 ω_2）的可能性是 1/2，因为知道袋子中白球数和黑球数都是 5 个．现引入一个定义如下：

定义 1　随机事件 A 发生可能性大小的度量（数值）称为 A 发生的概率，记作 $P(A)$.

对于一个随机事件而言，它发生可能性大小的度量是由它自身决定的，并且是客观存在的．一个根本的问题是，对一个给定的随机事件，它发生可能性大小的度量——概率究竟是多大呢？在前面的例子中，由于知道了袋子中白球和黑球数都是 5 个，才得以断定 $P(\omega_1)=1/2$. 如果不知道呢？但从实践经验可知，如果反复多次地从代中取球（取回放回），随着试验次数 n 的增大，比值 $\frac{n_{\text{白}}}{n}$ 会逐渐稳定到 1/2，记

$$\frac{n_{\text{白}}}{n}=\frac{\text{出现 }\omega_1\text{ 的次数}}{\text{试验总次数}}=f_n(\omega_1).$$

称 $f_n(\omega_1)$ 为事件 ω_1 在 n 次试验中出现的频率. 频率也在一定程度上反映了 ω_1 发生可能性的大小。尽管每作一串（n 次）试验，所得到的频率 $f_n(\omega_1)$ 可能各不相同，但只要 n 相当大，$f_n(\omega_1)$ 与 $P(\omega_1)$ 是会非常“靠近”的．因此，概率可以通过频率来测定，或者说频率是概率的一个近似．

下面较仔细地考察一下频率．如果随机事件 A 在 n 次反复试验中发生了 n_A 次，称

$$f_n(A) = \frac{n_A}{n}$$

为 A 的频率．易知频率具有下述性质：

(1)非负性：即 $f_n(A) \geqslant 0$；

(2)规范性：即若 Ω 为必然事件，则 $f_n(\Omega)=1$；

(3)有限可加性：即若 A、B 互不相容(即 $AB=\phi$)，则

$$f_n(A \cup B) = f_n(A) + f_n(B).$$

频率还具有一些别的性质，但是这三条性质是最基本的，其他的性质可以由它们推出．

因为频率的本质就是概率，因而有理由要求频率的这些性质也是概率所具有的．因为每一个随机事件 A，都有一个概率 $P(A)$ 与之对应，所以概率实质上是在事件域 $\mathscr{F}$ 上有定义的一个集合函数．它应该有下述性质：

(1)非负性：$P(A) \geqslant 0$ 对 $A \in \mathscr{F}$；

(2)规范性：$P(\Omega)=1$；

(3)有限可加性；若 $A_i \in \mathscr{F}, i=1,2,\cdots n$，且 $A_iA_j=\phi(i \neq j)$，则

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

这样一来，对于随机试验这样一个直观对象，就可以用“数学化”的语言来描述它们了．

第四节　古典概型

前面已经提到，一个随机试验，数学上是用样本空间 Ω，

事件域$\mathscr{F}$和概率P来描述的．对一个随机事件$A\in\mathscr{F}$，如何寻找它的概率$P(A)$是概率论的一个基本课题．先从最简单的随机试验入手．这类随机试验具有下述特征：

(1)样本空间的元素(即基本事件)只有有限个，不妨设为n个，并记为$\omega_1,\omega_2,\cdots,\omega_n$.

(2)每个基本事件出现的可能性是相等的，即有

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)$$

这种等可能的数学模型曾经是概率论发展初期的主要研究对象，常称这种数学模型为古典概型．它在概率论中有很重要的地位，一方面，因为它比较简单，许多概念既直观而易理解，另一方面，它又概括了许多实际问题，有很广泛的应用．

对上述的古典概型，它的样本空间$\Omega=\{\omega_1,\omega_w,\cdots,\omega_n\}$，事件域$\mathscr{F}$为$\Omega$的所有子集的全体．这时，连同$\phi$、$\Omega$在内，$\mathscr{F}$中含有$2^n$个事件，并且从概率的有限可加性知

$$1=P(\Omega)=P(\omega_1)+P(\omega_2)+\cdots+P(\omega_n).$$

于是 $P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)=\dfrac{1}{n}$.

对任一随机事件$A\in\mathscr{F}$，如果A是k个基本事件的和，则

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

$$=\frac{A\text{ 的有利事件数}}{\text{基本事件总数}}$$

A中所含的基本事件数，习惯上称为A的有利事件数，不难验证，上述概率$P(\cdot)$的确具有非负性，规范性和有限可加性.

现在举一些求$P(A)$的例子．

例1 在袋子中有10个相同的球，分别标为号码1，2，…，10，从中任取一个球，求此球的号码为偶数的概率．

解 令 $i=\{$所取球的号码为 $i\}$，$i=1,2,\cdots 10$，则

$$\Omega=\{1,2,3,\cdots,10\},$$

故基本事件总数为 $n=10$，又令

$$A=\{\text{所取球的号码为偶数}\}$$

显然 $$A=\{2\}\cup\{4\}\cup\{6\}\cup\{8\}\cup\{10\}$$

所以 A 中含有 $n_A=5$ 个基本事件，从而

$$P(A)=\frac{n_A}{n}=\frac{5}{10}=\frac{1}{2}.$$

例 2 一部四本头的文集按任意次序放到书架上去，问各册自右向左或自左向右恰为 1，2，3，4 的顺序的概率是多少？

解 若以 a、b、c、d 分别表示自左向右排列的书的卷号，则上述文集放置的方式可与向量 (a,b,c,d) 建立一一对应，因为 a、b、c、d 取值于 1，2，3，4，因此这种向量的总数相当于 4 个元素的全排列 $4!=24$. 又由于文集按"任意的"次序放到书架上去，因此这 24 种排列中出现任意一种的可能性相同，这是古典概型概率，其有利场合有 2 种，即自左向右或自右向左成 1、2、3、4 顺序，因此所求概率为 $\frac{2}{24}=\frac{1}{12}$.

例 3 有 10 个电阻，其电阻值分别为 $1,2,\cdots,10$，从中取出 3 个，要求取出的 3 个电阻，一个小于 5，一个等于 5，一个大于 5. 问取一次就能达到要求的概率是多少？

解 把从 10 个电阻中取出 3 个电阻的各种可能取法作为样本点全体，这是古典概率，其总数为 $\binom{10}{3}$，有利场合为 $\binom{4}{1}\binom{1}{1}\binom{5}{1}$，故所求概率为

$$P=\frac{\binom{4}{1}\binom{1}{1}\binom{5}{1}}{\binom{10}{3}}=\frac{1}{6}.$$

由以上的例题可以看出，求解古典概型问题的关键是寻求基本事件总数和有利事件数，但正面求这两个数并不是很容易的，有时要讲究一些技巧．

第五节 几何概率

在古典概型中利用等可能性的概念，成功地计算了某一类问题的概率. 不过，古典概型要求基本事件总数必须有限. 因此历史上有不少人企图把这种方法推广到有无限多结果而又有某种等可能性的场合．这类问题可以通过几何方法来解决．

先看几个例子：

例 1 某人午觉醒来，发现表停了，他打开收音机，想听电台报时．求他等待的时间短于 10min 的概率．

例 2 如果在一个 $50000\mathrm{km}^2$ 的海域里有表面积达 $40\mathrm{km}^2$ 的大陆架有石油，假如在这海域里随意选定一点钻探，问钻到石油的概率是多少？

一种相当自然的答案是认为例 1 所求的概率等于 1/6，例 2 中的概率等于 1/1250. 在求这些概率时，事实上是利用了几何的方法，并假定了某种等可能性．

在例 1 中，因为电台每小时报时一次，自然认为这个人打开收音机时处于两次报时之间，例如(13：00，14：00)，而且取各点的可能性一样，要遇到等待时间短于 10min，只有当他