概率论与数理统计

全国工科院校应用概率统计委员会
概率统计教材编写组编

上海科学技术出版社

概 率 论 与 数 理 统 计

全国工科院校应用概率统计委员会
概 率 统 计 教 材 编 写 组 编

上 海 科 学 技 术 出 版 社

责任编辑　唐仲华

概率论与数理统计

全国工科院校应用概率统计委员会
概率统计教材编写组编

上海科学技术出版社出版
(上海瑞金二路450号)
新华书店上海发行所发行　上海市印刷六厂印刷
开本 787×1092　1/32　印张 9　字数 194,000
1991年6月第1版　1991年6月第1次印刷
印数　1—15,700
ISBN 7-5253-2409-5/O·149
定价：3.60元

内容提要

本书为工科院校《概率论与数理统计》通用教材. 前五章除个别内容外，为必修部分，适合只讲授《概率论》的专业，附有习题答案. 第六、七两章为数理统计基本内容；方差分析与正交试验设计这一章作为数理统计的进一步应用.

各校可根据教学时数与专业特点有选择地讲授后三章内容. 由于用一个实测数据的例子贯穿全书的有关章节，便于教学时数较少的专业学生也有可能以自学形式学习后面的内容. 以达到能初步分析与处理随机数据的能力. 对计算较繁的内容，也提供部分计算程序供应用者参考.

为了便于教学，将为使用本教材的教师提供一本教学参考书，书中附有全部习题答案.

前　言

全国工科院校应用概率统计委员会于1989年4月在洛阳召开了《工科院校概率统计教学研究与应用成果学术交流会》. 会议对国内高等工科院校概率统计课程的使用情况展开了热烈的讨论，并对教材的进一步建设提出了许多有益的见解. 根据该会议关于教材建设的有关精神和与会同志的要求，委员会组织部分院校，成立概率统计教材编写组. 并于1989年6月在合肥工业大学召开了编写工作会议. 在广泛交流各校的教学经验基础上，共同拟订大纲和编写原则. 并于1990年1月，对写出的初稿，再次于合肥工大进行讨论与审定.

本书内容包括概率论及数理统计两部分. 为做到两者并重和力求有机结合，书中用一个样本数据的例子(§1.1中的例5)贯穿全书的有关章节. 力求做到理论概念有直观的统计背景，具体数据的统计分析与推断有理论根据. 并对较繁的公式计算及统计数据的处理方法提供相应的计算程序.

各章配有一定数量的习题，书末附有部分习题答案.

根据高等工科院校本科生学制特点和专业要求,本书除选取了一般公认的内容外,在选材上更注重各种概率统计方法的实用价值. 在编写方法上力求纲目清晰，突出基本概念、基本理论方法和基本计算. 为了能在计算机上实际使用各种统计方法，本书在附录中提供了部分计算程序供需要的

教师和读者参考使用．程序以书中所列举例子的数据计算其结果．

参加本书编写的作者有青岛建筑工程学院陈鸣皋、西安石油学院茹世才、河北煤炭建筑工程学院王清印、沈阳航空工业学院吴晓明、四川建筑材料工业学院舒启国、哈尔滨建筑工程学院杨有贵、合肥工业大学戴俭华和福州大学郭福星、林可容．并由郭福星、戴俭华担任主编．

参加本书审稿的同志有青岛建筑工程学院安世仁、四川建筑材料工业学院舒启国、安徽建筑工业学院张福仁．

侯振挺担任本书的主审．全国工科院校应用概率统计委员会主任侯振挺教授热情关怀本书的编写与出版工作，他在百忙中仔细审阅了全书的初稿．

本书的全部编写工作是在全国工科院校应用概率统计委员会交流培训部直接主持下进行的．参加编写与审稿的上述院校给予了大力支持和帮助．此外，本书能顺利地同读者见面，还由于上海科学技术出版社的大力支持，编者向他们一并致谢．

由于水平所限，加之时间仓促，谬误之处恐所难免．希望使用本书的同行和广大读者批评指正．

编　者

1990年12月

目　　录

第一章　随机事件与概率

概率论与数理统计是一门研究和揭示随机现象的统计规律性的学科. **所谓随机现象，是指在同一条件下，所观察的现象(或实验结果)可能发生，也可能不发生.** 譬如:

例 1　向上垂直投掷一枚质地均匀的硬币，其落地后是“国徽”(正面)或“文字”(反面)朝上，事先无法预言;

例 2　射手在一次打靶(标有 0～10 环)中，“击中第 10 环”这一现象可能发生，也可能不发生;

例 3　一张有奖储蓄券(每一万户设 1～4 等奖各为 2、10、100、1000 名)，在未开奖之前，无法预言是否中奖;

例 4　从一批装有 100 个产品，且含有 5%次品的产品中任意抽验一件产品，其结果可能是正品，也可能是次品;

例 5　自动装包机装包某类产品的重量标准为 100kg. 现从成品中任取一包测量，其“重量＜100kg”可能发生，也可能不发生;

……

我们对上述某一现象所作的一次观察或进行的一次实验，统称为**随机试验**，或简称为**试验**，常用记号 E 表示. 它们具有如下特点:

1. 试验可以在相同条件下重复进行;

2. 每次试验可能出现的结果不只一个，但在试验之前无法预言究竟出现哪一个.

在随机试验中，可能发生也可能不发生的现象称为随机

事件，或简称为事件，通常用大写字母 A、B、C、……表示.

把试验中不能再分解的事件称为基本事件. 譬如例 1 的基本事件有 2 个；例 4 的基本事件为 100 个.

因此，上面所述各例中的现象都是随机事件，所不同的是它们在试验后出现的可能性有的较大，有的则较小. 譬如例 3，若某人持有一张有奖储蓄券，那么其在开奖后，不中奖的可能性远比中奖的可能性大得多. 而对例 1 而言，硬币落地后正、反面朝上的可能性应该是一样的. 这些结论目前只能是从经验中获得.

在客观世界里，随机现象是普遍存在的. 有些部门，譬如气象、地震、水文等，都是有针对性地把随机事件作为自己的研究对象的，从而产生了预测或预报的技术. 同随机事件有本质差别的是肯定性事件，它指的是在每次试验中必然要发生的现象(称为必然事件)，或不可能发生的现象(称为不可能事件)这两种极端情况. 前者常记为 Ω，后者记为 Φ. 譬如“北京 7 月份的平均气温 > 10℃”就是一必然事件;而“北京 1 月份的平均气温 > 30℃”则为一不可能事件.

为今后研究的方便起见，也把 Ω、Φ 当作(特殊的)随机事件. 读者将在 § 1.2 看到这种规定的必要性.

本章的目的在于揭示随机事件的统计规律性的意义，并对事件出现的可能性大小提供明确的计算方法.

§ 1.1 概率的直观意义

怎样用数值来确切地表示随机事件 A 出现的可能性大小呢？在概率论发展的初期，人们在长期的实践中发现，对于随机事件 A，若在 n 次试验中出现了 m 次，则 n 次试验中 A 出现的频率

$$f_n(A) = m / n ,$$

在n较大时就会呈现出明显的规律性.

例 1 历史上有多位著名科学家，曾做过成千上万次的抛掷硬币的试验，并统计了 n 次试验中出现正面（事件 A）的次数 m，及相应的频率 $f_n(A) = \frac{m}{n}$ 如下表所示.

表 1.1.1 抛掷一枚硬币的试验结果

实验者	掷币次数 n	出现正面次数 m	频率$f_n(A)$
狄莫弗	2048	1061	0.5181
蒲丰	4040	2048	0.5069
弗勒	10000	4979	0.4979
皮尔逊	24000	12012	0.5005

从上述结果可以看出：当试验次数 n 较大时，比值 $f_n(A) = m / n$ 总是围绕在 0.5 附近摆动.

这个稳定值说明随机事件 A 发生的可能性大小是客观存在的，是不以人们的意志为转移的客观规律，我们把它称为随机现象的**统计规律性**．读者自己也可以做这个试

验，当你抛掷的试验次数 n 较小时，事件 A 的频率 $f_n(A) = m/n$ 是不稳定的.

如果用一个数来表示这种大小，可得如下概念.

定义 1(概率的统计定义) 设在相同条件下对 E 重复进行 n 次试验，其中事件 A 出现 m 次．当试验次数 n 充分大时，事件 A 出现的频率 $f_n(A)=\frac{m}{n}$ 的稳定值 p，称为事件 A 的概率，记为 $P(A)$. 即

$$p=P(A)\approx f_n(A)=\frac{m}{n} \qquad (m\leqslant n).$$

因为必然事件的频率恒等于 1，不可能事件的频率恒等于 0，所以由频率的性质可以推想概率至少具有如下基本性质:

1. **非负性**: 对任意的随机事件A，有

$$0\leqslant P(A)\leqslant 1. \tag{1.1.1}$$

2. **规范性**: $P(\Omega)=1(P(\Phi)=0)$. (1.1.2)

在实际应用中，当重复试验的次数充分大时，可用事件的频率作为其概率的近似值．

例 2 多次抽查某工厂的产品，每次抽查 n 件，统计出现合格品(事件 A)的件数 m，及其相应合格率(即合格品出现的频率 $f_n(A)$)的结果如下表．

表 1.1.2 合格率的稳定性试验

抽查数 n	10	60	150	600	900	1200	1800	2400
合格数 m	7	53	131	548	820	1091	1631	2152
合格率$\frac{m}{n}$	0.7	0.833	0.873	0.913	0.911	0.909	0.906	0.897

与例 1 类似，虽然在有限的 n 次抽查产品时，抽到合

格品（事件 A）的次数 m 具有随机性，但随着抽查件数 n 的增加，比值 $f_n(A)=\dfrac{m}{n}$ 总是在常数 0.9 附近摆动．

对于某些特殊类型的随机试验，譬如古典型的与几何型（见 §1.3）的随机试验，其事件发生的频率的稳定值可根据下面将要定义的古典概型和几何概型直接求出来．

古典型随机试验是指具有下列两个特征的随机试验：

1. **有限性**：试验的所有可能结果只有有限个基本事件；

2. **等可能性**：每次试验中各可能结果出现的可能性均相同．

古典型随机试验也叫古典概型试验，简称之为古典概型．例如抛硬币试验，由于试验只有两种可能结果："正面朝上"（事件 A）或"反面朝上"（事件 B），而且在每次试验中事件 A 或事件 B 的出现都是等可能的，所以这种试验是古典概型的．

定义 2（概率的古典定义） 设一个试验只有 n 个等可能的基本事件，其中 m 个基本事件具有属性 A，则称 m/n 为事件 A 的概率．记为

$$P(A)=\frac{\text{具有属性}A\text{的基本事件数}}{\text{所有可能的基本事件数}}=\frac{m}{n}. \tag{1.1.3}$$

例如，在抛硬币试验中，每次试验只有两种基本事件："正面朝上"（事件 A）与"反面朝上"（事件 B）．且出现 A 与 B 的可能性相同（设硬币均匀对称）．于是按式（1.1.3），有 $P(A)=\dfrac{1}{2}$．

怎样判断事件出现的"等可能性"？这要从事件是否

完全处于平等的地位来考虑．譬如前言中的例 4，每个产品在一次的任取中被取到的机会是均等的，所以这 100 个产品的每一个都是等可能的基本事件．以 A 表示取到次品，B 表示取到正品，则有 $P(A)=\frac{5}{100}$，$P(B)=\frac{95}{100}$．

显然，前言中的例 2 就不能用式 (1.1.3) 计算，虽然打靶的结果是有限个事件，但其打中各环的机会是不均等的．故此类问题只能由重复试验的统计规律性来给出射手击中目标的概率．

例 3　设袋中有 10 个外形相同的有色球 (6 红 4 白)，现从中任取 3 个，试求抽到 3 个都是红球 (事件 A) 及抽到 2 个红球、1 个白球 (事件 B) 的概率各为多少?

解　从 10 个球中随机地抽取 3 个，共有 C_{10}^3 种等可能的取法，每种取法对应着一个基本事件，所以基本事件的总数为 $n=C_{10}^3=120$．具有属性 A 的基本事件数为 $m=C_6^3=20$．所以

$$P(A)=20/120=1/6.$$

因为具有属性B的基本事件数为$m=C_6^2\cdot C_4^1=60$，所以

$$P(B)=60/120=1/2.$$

本例中的“从中任取 3 个”，如无特别声明怎样取法，都理解为每次取出 3 个．此类模型的例子广泛应用于实际问题的计算．产品抽样检查就是其中之一．

抽样是指从待查的整批产品中抽取部分产品．在数理统计中，该部分产品叫做**样本**或**子样**，样本中所包含的产品数叫做样本的**容量**，而待查的整批产品叫做**总体**或**母体**，随机抽样指的是从总体中的每一件产品都有同等的可

能性被取做样本中的样品.

例 4 设有一批产品共 N 件，其中有 M 件次品. 现在从全部 N 件产品中随机地抽取 n 件 $(n \leqslant N)$，试求恰好取到 m 件 $(m \leqslant M)$ 次品的概率.

解 设 A 表示“n 件产品中有 m 件次品”这一事件，从 N 件产品中抽取 n 件产品的所有等可能的基本事件数为 C_N^n 种. 事件 A 的发生相当于从 M 件次品中抽取 m 件次品以及从 $N-M$ 件正品中抽取 $n-m$ 件正品，故具有属性 A 的基本事件数为 $C_M^m C_{N-M}^{n-m}$. 因此

$$P(A)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n} . \tag{1.1.4}$$

程序Pro1 提供 C_n^k 及上式的计算①.

一个工厂生产的产品，其废品率严格地说是未知的，因为它的全部产品构成一个总体. 一般说来，我们无法对这个总体的每一个个体(特别是一些具有破坏性的产品，如电子元件等)进行抽查，而只能从中抽取一个数量为 n 的样本进行检查，以便从这个容量为 n 的样本中的随机结果去推断总体的情况. 关于这个课题的研究构成数理统计的重要内容.

由于抽样具有随机性，因而对于不同的抽样可能得到不同的结果. 这就迫使我们对各种结果出现的可能性大小进行讨论. 下面列举的例子将贯穿于本书的有关章节，以此体现我们所讨论的这些章节的研究对象及其研究的方法等,请读

①本书对于一些计算较繁的公式或数据处理的方法给出相应的计算程序,其使用方法在附录中有统一的说明，目的在于为使用者提供方便，具有相对的独立性. 对于暂时不用的读者可以略去，它不属于本书的基本要求.

者注意.

例 5 为了解自动装包机的工作是否正常(装包的标准重量 $W=100\text{kg}$),今从某日生产的产品中抽查 130 包,具体数据如下:

101.1	100.6	101.1	101.7	102.4	102.7	103.2	103.7
99.6	99.1	98.6	98.1	97.6	97.1	96.8	97.3
97.7	97.8	98.2	98.3	98.3	98.4	103.1	102.8
102.6	102.5	102.3	101.9	101.2	101.1	99.6	99.7
99.9	99.9	99.1	98.6	102.2	102.1	102.3	101.8
101.7	101.6	102.0	101.8	101.8	102.0	101.5	101.3
101.4	100.9	101.0	100.6	98.6	101.0	100.9	100.8
100.7	100.6	101.4	101.3	101.5	101.4	101.3	101.2
100.7	99.0	100.9	101.0	100.7	100.8	99.2	99.3
99.4	99.5	99.5	99.4	99.3	99.2	99.1	99.6
100.6	101.0	100.9	100.2	100.3	100.4	100.5	100.3
100.0	100.0	99.9	99.8	99.7	99.6	100.1	100.2
100.3	100.4	100.5	100.3	99.6	99.7	99.8	99.6
100.0	100.0	100.3	100.5	100.4	100.2	100.1	100.2
100.3	100.4	100.5	99.0	98.7	98.8	98.9	98.7
99.0	98.6	99.5	99.4	99.3	99.1	99.7	100.1
100.2	100.3						

若规定"重量 $|W-100|>3$"(记为 A)者为不合格品,试求从这样本中任取一包为不合格品的概率为多少?

解 经检验知道有 4 包是不合格的,所以若从这个样本中任取一包为不合格品的概率是

$$P(A)=\frac{4}{130}=0.031.$$