大　学　数　学　教　程

概率论与数理统计

第二版

山东大学数学学院
刘建亚　吴　臻　主编
胡发胜　叶　宏　吕　同　编

高等教育出版社
HIGHER EDUCATION PRESS

大学数学教程

概率论与数理统计

Gailülun yu Shuli Tongji

第二版

山东大学数学学院

刘建亚　吴臻　主编

胡发胜　叶宏　吕同　编

高等教育出版社·北京

HIGHER EDUCATION PRESS　BEIJING

内容简介

本书内容少而精，突出数学思想，注重理论与应用相结合，内容与中学知识相衔接，深入浅出，易教易学. 本书主要内容包括随机事件及其概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、数理统计的基本知识、参数估计和假设检验、一元线性回归分析和方差分析.

书中每章最后一节均是解决本章主要问题的 MATLAB 程序和例题演示. 每章后的习题配置基本练习题和部分综合练习题，书末附有习题答案. 附录给出了概率统计发展简介以及概率统计在数学建模中的应用.

本书可供高等学校非数学类专业学生使用，也可供科技工作者学习参考.

图书在版编目(CIP)数据

大学数学教程. 概率论与数理统计/刘建亚，吴臻主编；胡发胜，叶宏，吕同编. —2版. —北京：高等教育出版社，2011.7

ISBN 978-7-04-032301-6

Ⅰ.①大… Ⅱ.①刘…②吴…③胡…④叶…⑤吕… Ⅲ.①高等数学-高等学校-教材②概率论-高等学校-教材③数理统计-高等学校-教材 Ⅳ.①O13②O21

中国版本图书馆 CIP 数据核字(2011)第121557号

策划编辑 于丽娜　责任编辑 马丽 李蕊　封面设计 张志奇　版式设计 马敬茹
插图绘制 黄建英　责任校对 刘春萍　责任印制 张泽业

出版发行	高等教育出版社	网　　址	http://www.hep.edu.cn
社　　址	北京市西城区德外大街4号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.landraco.com
印　　刷	中国农业出版社印刷厂		http://www.landraco.com.cn
开　　本	787×960　1/16		
印　　张	14.5	版　　次	2003年7月第1版
字　　数	270 000		2011年7月第2版
购书热线	010-58581118	印　　次	2011年7月第1次印刷
咨询电话	400-810-0598	定　　价	20.20元

物 料 号　32301-00

第二版前言

本套教材是由山东大学数学学院具有丰富教学经验的一线教师编写的，第一版是普通高等教育“十五”国家级规划教材，包括《微积分 1》、《微积分 2》、《线性代数》、《概率论与数理统计》、《复变函数与积分变换》五册．经过多年的教学实践，山东大学数学学院在大学数学课程建设和教学改革方面取得了可喜的成绩．微积分与数学实验、线性代数、复变函数与积分变换分别于 2007 年、2006 年、2010 年被评为国家精品课程，以刘建亚教授为带头人的大学数学系列课程教学团队被评为 2007 年度国家级教学团队．

为更好地将优秀教学改革成果运用并推广，根据当前的教学实际，山东大学数学学院组织中坚力量，对第一版进行了修订．在保持本书第一版优点、特色的前提下，新版教材注重与中学教学内容的衔接，增加了与中学数学接轨的部分内容；增选一些国外教材中的案例、例题和习题，力求题型新颖．为更好地将数学建模思想融入教学，培养学生的建模思想和意识，通过增设有关章节介绍与教学内容相关的建模案例，全方位提升学生的综合素质和创新能力．新版教材力求做到符合大学数学课程教学基本要求，知识结构符合认知规律，同时渗透现代数学思想，加强应用能力培养，便于学生学习和教师教学．

本书为《概率论与数理统计》分册，与第一版相比，新版主要作了如下修改：

1. 根据教学实际，增加了有关基本概念、基本运算的例题，使之更加符合学生的学习和思维习惯．

2. 对原有的习题作了调整，扩大了题目涉及的范围，增加了题型的变化，另外，补充了许多考研经典题目，以满足学生考研复习的需要．

3. 对于各部分内容作了补充与完善，例如，关于多维随机变量函数的分布，增加了离散型随机变量函数的情形；中心极限定理中增加了棣莫弗－拉普拉斯定理等．另外，对于很多定理及方法，补充了证明或者分析，使之更加实用化和通俗化，便于学生理解和使用．

4. 为适应当今社会对数学建模的需要，在附录增加了与概率统计内容对应的数学建模应用实例．

5. 将第一版附录中的著名数学家简介修改为以概率论与数理统计的发展简史为主线、穿插数学家的简介．

本次修订工作主要由吕同(第1章)、叶宏(第2,3,4章)、胡发胜(第5,6,7章)完成. 刘建亚教授、吴臻教授按照丛书总体要求对修改框架提出具体建议,栾贻会教授对全书进行了审定. 另外,数学实验内容由傅国华编写,概率论与数理统计发展简史由包芳勋编写. 在本书修订过程中,我们得到高等教育出版社、山东大学教务处、山东大学数学学院领导及同事的大力支持,兄弟院校的同行也对此次修订提出了不少宝贵建议. 在此,我们表示衷心的感谢.

限于编者水平所限,新版中难免存在不足,欢迎广大专家、同行与读者批评指正.

编　者

2011年2月

第一版前言

为了适应新世纪我国高等教育迅速发展的形势和实行学分制的需要，满足新时期高等教育人才培养拓宽口径、增强适应性对数学教育的要求，山东大学数学与系统科学学院从2000年开始按照教育部《高等教育面向21世纪教学内容和课程体系改革计划》的精神和要求，在学院领导的亲自参与下，组织部分教师对非数学类专业大学数学的课程体系进行了认真深入的研究和论证．针对大学数学是高校非数学类专业所有大学生应当具有的素质，又考虑到不同专业的要求深浅不同、内容多少各异的实际情况，制订了适应这种情况的新课程体系．新课程体系的主要特点是采取平台加模块的结构，整个大学数学的课程共分三个平台，不同平台反映了不同专业对数学知识的不同层次、级别要求，体现数学知识结构和大学生认知结构的统一．鉴于人类认识是从感性到理性，由易到难，由浅入深的，因此第一平台(包括微积分(一)、线性代数和概率统计)是体现高等数学的普及和基础，体现所有各专业应当具有的数学素质教育，主要侧重基本概念和基本方法，加强基本运算，努力渗透基本数学思想；第二平台是对第一平台基本概念的加深和认识方法的拓宽，在本平台中还适当体现出数学理论的系统性和严谨性；第三平台(包括数学建模、数值分析、数理方程、复变函数和积分变换、运筹学等)则是为满足某些对数学知识和方法有特殊要求的专业而设置．各平台的教学内容由浅入深，反映不同专业对数学知识和内容的不同要求；各平台的内容又采取模块组合的方式，模块间相对独立，各专业亦可根据本专业的需要，选用不同的模块组合，这样就使得新的课程体系具有更大的灵活性，能够满足不同层次、不同要求的专业对数学教学的需求．另外，新课程体系还将利用计算机解决数学问题的数学实验融入其中，做到理论和实践的有机结合．

山东大学教务处对新课程体系给予充分的肯定，并大力支持按新课程体系编写相应的教材．在我们完成初稿之后，教务处安排几个专业的学生先行试用，并在此基础上加以修改完善．目前，已完成了前两个平台共计四册的教材编写和修改．其中，微积分为两册，分属两个平台；线性代数和概率统计各一册．其中《概率论与数理统计》这本教材还有以下特色：

1．内容少而精，体现素质教育，突出数学思想．我们重点介绍概率论与数理统计中的基本概念和基本方法；从培养能力和提高素质为着眼点，有选择地保留了部分定理、性质的证明，对那些用类似的技巧方法，或者读者举一反

三可以理解或自学的证明部分作省略或简化处理.

2. 扩大了读者的知识面. 我们将各专业不同需求的数学内容融进了一套教材中. 主要的做法是: 用“*”号标明不同层次对数学的要求; 从不同的学科例题分析中引进基本概念; 习题中也涉及多学科. 如在数学要求较低专业学习的读者希望学习更多数学知识(如跨学科考研或工作需要)时, 可以从同一本书按“*”号的标示获取. 当然, 教师在授课时可按本专业的要求有选择地使用.

3. 各节后的习题配置除基本练习外, 还有部分综合练习题, 以提高读者分析问题、解决问题的能力. 综合练习题多置于每节习题后且配以“*”号标示.

4. 紧密联系实际问题, 适当反映了统计方法在实际应用中的新进展. 增添了利用计算机解决数学问题的内容, 在每章后均有解决本章主题问题的 MATLAB 程序和例题演示.

本套书由山东大学数学与系统科学学院组织部分有较高水平和丰富教学经验的教师集体编写, 最后聘请有关专家审定. 在长达近两年的编写过程中, 学院领导给予了极大的关注、支持和具体指导, 为此曾多次召开各种类型的会议反复论证, 几易手稿.

大学数学教程的主编是刘建亚, 概率统计部分由吕同(第1、2、3、4章)、胡发胜(第5、6、7章)编写, 由胡发胜完成修改及统稿工作; 吴臻、刘锦萼审查定稿; 各册的数学实验内容由傅国华编写和制作.

本套教材作为普通高等教育“十五”国家级规划教材正式出版, 是教育改革的产物. 在此, 我们感谢山东大学教务处、山东大学出版基金委、山东大学数学学院领导对改革和教材出版的鼎力支持, 感谢仪洪勋、江守礼教授对我们的鼓励和帮助. 我们特别感谢高等教育出版社, 由于他们的指导和帮助才使本书顺利与读者见面.

新时期大学数学的教学改革是一项非常紧迫, 非常重要, 也是非常艰巨的工作. 限于编者水平, 本书肯定会有许多不足和缺点乃至问题, 恳请读者批评指正.

编　者

2002 年 11 月

目　　录

第1章　随机事件及其概率

自然界和人类社会中出现的种种现象，大体上可分为两类：一类现象是在一定条件下必然发生或绝不可能发生的，此类现象称为**确定性现象**. 例如，在1个标准大气压下把水加热到100 ℃时必然会沸腾；电流通过导线时，导线周围必然产生磁场，等等. 另一类现象则不然，即在相同的条件下可能发生也可能不发生，或者说，可能出现这个结果也可能出现那个结果，呈现出偶然性，此类现象称为**随机现象**. 例如，抛掷一枚硬币究竟是正面(币值面)朝上还是反面朝上，在每次抛掷之前是无法判定的；观察用一种新药治疗某种疾病的疗效，对一个病人来说，可能有效也可能无效. 通过大量试验我们知道：大量重复抛掷同一枚硬币时正面朝上的次数约占抛掷总数的一半. 这类在个别观察试验中呈现不确定的结果，而在相同条件下，大量重复试验，试验结果呈现出的规律性称之为统计规律性. 又如，为评价一种新药的疗效，通过足够多个病例的试用和观察，可以对其效果作出客观的估计. 概率论与数理统计就是研究随机现象统计规律性的一门数学学科，它在自然科学、工程技术和社会科学的许多领域中有着重要的应用. 特别是随着计算机的普及使用，概率统计在经济管理、金融保险、生物医药等方面的应用更加广泛、更加深入.

§1.1　随机事件及其运算

1. 随机试验与样本空间

为了确定随机现象的规律性，需要进行多次的试验、实验、调查或观察，我们把这些工作统称为试验. 概率论中所说的试验是指随机试验，它具有下列三个特征：

1) 试验可在相同的条件下重复进行；

2) 试验的结果不止一个；

3) 每次试验之前，不能判定哪一个结果将会出现.

例如，掷一颗骰子，观察出现的点数；记录某传呼台在一小时内接收到的呼唤次数；观察日光灯的使用寿命，以及前面提到的掷硬币试验等，都是随机试验，简称为试验，用 E 表示.

试验 E 中的每一个可能结果称为基本事件，或称为样本点，所有基本事件组成的集合称为试验 E 的样本空间，记为 Ω.

例 1.1.1 在抛掷一枚硬币试验中，有两个可能的结果：出现正面，出现反面. 若分别用“正”、“反”来表示，即有两个基本事件，这个试验的样本空间为 $\Omega=\{正,反\}$.

例 1.1.2 掷一颗骰子，观察其出现的点数，所有可能出现的结果有 6 个：出现 1 点，出现 2 点，…，出现 6 点. 分别用 1,2,…,6 表示，则样本空间为 $\Omega=\{1,2,\cdots,6\}$.

例 1.1.3 在一批日光灯中任意抽取一只，测试其寿命，用 t(单位:小时)表示日光灯的使用寿命，则 t 可取所有非负实数：$t\geqslant 0$，对应了试验的所有可能结果，则样本空间为 $\Omega=\{t\mid t\geqslant 0\}$.

由上述例题可以看出，样本空间应根据随机试验的内容来确定，这是很重要的.

2. 随机事件

随机试验 E 的样本空间 Ω 中的子集称为试验 E 的**随机事件**，简称为**事件**，常用大写英文字母 A，B，C，…表示. 在例 1.1.2 中，基本事件{出现 2 点}、{出现 4 点}、{出现 6 点}以及由它们组成的集合“出现偶数点”，都是该试验的随机事件. 基本事件是最简单的随机事件，而一般的随机事件是由若干个基本事件组成的，称为**复合事件**.

随机事件中有两个极端的情况：一是由样本空间 Ω 中的所有元素组成的集合，称之为必然事件，用 Ω 来表示，它在每一次试验中都发生. 例如，前面所述“抛掷一颗骰子，出现点数都不大于 6”就是必然事件. 另一种是不含任何元素的空集合，称之为不可能事件，用 $\varnothing$ 来表示，例如“抛掷一颗骰子，出现点数大于 6”就是不可能事件. 它在每一次试验中都不会发生. 严格来说这两种事件不是随机事件，但为了今后讨论方便，我们还是把必然事件与不可能事件作为随机事件的特殊情形来统一处理.

3. 事件间的关系与运算

在同一随机试验中，事件不止一个. 有些事件简单，有些事件复杂. 通过研究它们之间的联系. 可以更好地帮助我们理解事件的本质.

设试验 E 的样本空间为 Ω，A，B，C，$A_k(k=1,2,3,\cdots)$ 是 E 的事件.

(1) **包含** 若事件 A 发生必然导致事件 B 发生，则称事件 B **包含事件** A，或称**事件** A **包含于事件** B **中**，记为 $B\supset A$ 或 $A\subset B$，如图 1.1.1 所示(其中,事件 A 发生和 B 发生分别表示样本点落在圆 A 和圆 B 内,下同).

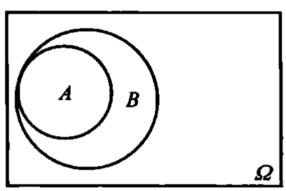

图 1.1.1

显然，必然事件 Ω 包含任何事件 A，事件 A 包含不可能事件 $\varnothing$，即 $\Omega\supset A\supset\varnothing$.

特别地，若事件 A 包含事件 B，且事件 B 也包

含事件 A，即 $A \supset B$ 且 $A \subset B$，则称事件 A 与事件 B **相等**，记为 $A=B$.

(2) **事件的并** 若事件 A 与事件 B 至少有一个发生，这样构成的事件称为事件 A 与事件 B 的**并事件**(或称为 A 与 B 的和事件)，记为 $A \cup B$. 例如，10 件产品中有 3 件次品，从中任取 2 件，若 A 表示"取到 1 件次品"，B 表示"取到 2 件次品"，则和事件 $A \cup B$ 表示"至少取到 1 件次品".

事件 $A \cup B$ 通常包含三个部分：A 发生而 B 不发生，A 不发生而 B 发生，A，B 都发生. 如图 1.1.2 阴影部分所示.

类似地，n 个事件 $A_1, A_2, \cdots, A_n$ 的并事件 $A_1 \cup A_2 \cup \cdots \cup A_n$ 表示"$A_1, A_2, \cdots, A_n$ 中至少有一个发生".

(3) **事件的交** 由事件 A 与事件 B 同时发生而构成的事件称为事件 A 与事件 B 的**交事件**(或称为 A 与 B 的积事件)，记为 AB(或 $A \cap B$). 如图 1.1.3 阴影部分所示.

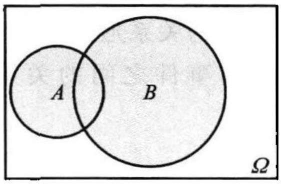

图 1.1.2

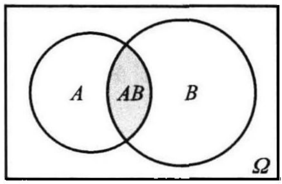

图 1.1.3

在上面"10 件产品中有 3 件次品"的例子中，$A \cap B = \varnothing$；若用 C 表示"至多取到 2 件次品"，则 AC 表示"恰好取到 1 件次品"，即 $AC=A$.

类似地，n 个事件 $A_1, A_2, \cdots, A_n$ 的交事件 $A_1 A_2 \cdots A_n$ 表示"$A_1, A_2, \cdots, A_n$ 同时发生".

(4) **事件的差** 事件 A 发生而事件 B 不发生，这样构成的事件称为事件 A 与事件 B 的**差事件**，记为 $A-B$. 如图 1.1.4 阴影部分所示.

(5) **互不相容事件** 在一次试验中，若事件 A 与事件 B 不能同时发生，即 $AB=\varnothing$，则称 A 与 B 为**互不相容事件**(或称为互斥事件). 例如，掷一颗骰子，A 表示"出现 3 点"，B 表示"出现 4 点"，则 A 与 B 为互不相容事件. 如图 1.1.5.

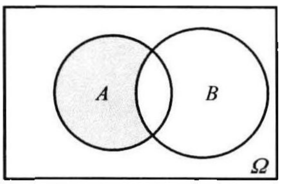

图 1.1.4

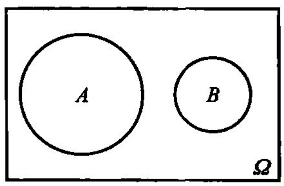

图 1.1.5

一般地，对于 n 个事件 $A_1, A_2, \cdots, A_n$，若它们之间两两互不相容，则称这 n 个事件是**互不相容**的. 通常 $A_1 \cup A_2 \cup \cdots \cup A_n$ 记为 $A_1 + A_2 + \cdots + A_n$ 或 $\sum_{i=1}^{n} A_i$.

（6）**对立事件**　在一次试验中，若事件 A 与事件 B 二者必有一个发生且仅有一个发生，则称 A 与 B 为**对立事件**（或称为互逆事件），通常把 A 的对立事件记为 $\overline{A}$（即 $B = \overline{A}$），$\overline{A}$ 也称为 A 的逆事件. 例如，掷一枚硬币，用 A 表示"出现图案面"，而事件 B 表示"出现币值面"，则 A 与 B 为对立事件. 如图 1.1.6 所示.

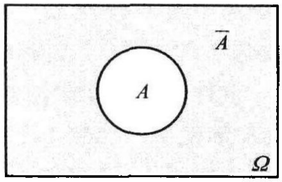

图 1.1.6

显然，$A\overline{A} = \varnothing$，且 $A \cup \overline{A} = \Omega$.

值得注意的是，若事件 A 与 B 对立，则事件 A 与事件 B 互不相容，反之不然.

概率论中事件间的关系和运算与集合论中集合间的关系形式上是类似的，利用在中学里学到的集合知识，可以更好地理解事件之间的关系（见表 1.1.1）.

表 1.1.1

符号	概率论中	集合论中
Ω	必然事件	全集
$\varnothing$	不可能事件	空集
A	事件	子集
$\overline{A}$	A 的对立事件	余集
$A \cup B$	事件的并	并集
$A \cap B$	事件的交	交集
$A - B$	事件的差	差集
$A \cap B = \varnothing$	互不相容事件	A 与 B 没有公共元素

集合论中常用的运算律同样适用于概率论的事件的运算.

（1）**交换律**　$A \cup B = B \cup A$,

$A \cap B = B \cap A$;

（2）**结合律**　$(A \cup B) \cup C = A \cup (B \cup C)$,

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

（3）**分配律**　$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

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

（4）**德摩根定律**　$\overline{A \cap B} = \overline{A} \cup \overline{B}$,

$$\overline{A\cup B}=\overline{A}\cap\overline{B}.$$

对于 n 个事件 $A_1,A_2,\cdots,A_n$ 有

$$\overline{A_1\cup A_2\cup\cdots\cup A_n}=\overline{A}_1\cap\overline{A}_2\cap\cdots\cap\overline{A}_n,$$

$$\overline{A_1\cap A_2\cap\cdots\cap A_n}=\overline{A}_1\cup\overline{A}_2\cup\cdots\cup\overline{A}_n.$$

例 1.1.4　设 A，B，C 表示任意三个随机事件，用 A，B，C 及其运算符号表示下列事件：

(1) A 发生而 B 与 C 都不发生；

(2) A 发生且 B 与 C 至少有一个发生；

(3) A 与 B 发生而 C 不发生；

(4) A，B，C 三个事件中只有一个发生；

(5) A，B，C 中至少有两个发生；

(6) A，B，C 中至多有一个发生.

解　(1) 该事件可表示为 $A\overline{B}\,\overline{C}$ 或 $A-B-C$ 或 $A-(B\cup C)$；

(2) 因为 $B\cup C$ 表示 B 与 C 至少有一个发生，故该事件可表示为 $A\cap(B\cup C)$；

(3) A 与 B 发生即 A 和 B 都发生，就是 AB，C 不发生即 $\overline{C}$ 发生，故该事件可表示为 $AB\overline{C}$ 或 $AB-C$；

(4) 该事件为 $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$；

(5) 该事件为 $AB\cup BC\cup CA$；

(6) 该事件为 $\overline{A}\,\overline{B}\cup\overline{B}\,\overline{C}\cup\overline{C}\,\overline{A}$.

§1.2　随机事件的概率

研究随机现象，我们不仅需要知道它可能出现哪些事件，更需要知道每个事件出现的可能性的大小. 所谓事件的**概率**，就是刻画事件出现的可能性大小的一种数量指标，这个数量指标应满足以下两个要求：

(1) 它应是事件本身固有的，不随人们的意志而改变的一种客观属性量度.

(2) 它必须符合一般常情，即事件发生可能性大的，它的值就大；事件发生可能性小的，它的值就小.

1. 频率与概率

定义 1.2.1　**设在相同的条件下，进行了 n 次试验，在这 n 次试验中事件 A 出现了 m 次，则称**

$$f_n(A)=\frac{m}{n}=\frac{\text{事件 }A\text{ 出现的次数}}{\text{试验的总次数}}\tag{1.2.1}$$

为随机事件 A 在 n 次试验中出现的频率，m 称为频数.

由频率的定义，容易得出它具有下列基本性质：

$$0 \leqslant f_n(A) \leqslant 1,\ f_n(\Omega)=1 \quad (f_n(\varnothing)=0).$$

若 $A_1, A_2, \cdots, A_n$ 是两两互不相容事件，则

$$f_n(A_1 \cup A_2 \cup \cdots \cup A_n) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_n).$$

事件的频率刻画了事件发生的频繁程度，由(1.2.1)式容易看出，频率越大，事件 A 发生越频繁，因而在一次试验中发生的可能性越大. 因此，我们自然会想到用事件 A 的频率表示 A 在一次试验中发生的可能性的大小. 但频率是随试验的具体结果而定的，一般不是确定的值.

应该指出，随机事件的频率与我们进行的试验有关，而随机事件的概率则是完全客观地存在着的，随机事件的频率 $f_n(A)$ 可以看作是它的概率的随机表现. 如：某人进行 n 次试验，A 发生的次数为 m 次. 另一人再进行 n 次相同的试验，则 A 发生的次数可能仍是 m 次，也可能不再是 m 次. 因此，频率总是对特定的 n 次试验而言的. 大量经验表明，当试验的次数相当大时，频率总是稳定于某一常数附近，即它将以某一常数为中心作微小的摆动，而发生较大偏离的可能性很小. 这一性质称为**频率的稳定性**. 下面的例子说明了这一点.

例 1.2.1　掷一枚质地均匀的硬币，出现正面与反面的机会是相等的，即在大量重复试验中，出现正面的频率应接近于 0.5. 历史上曾有几位数学家做过该试验，结果见表 1.2.1.

表 1.2.1

试验者	掷硬币次数	正面朝上次数	频　率
蒲　丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

例 1.2.2　瑞典 1935 年官方统计资料显示该年各个月份出生的男婴和女婴人数，以及生女的频率，见表 1.2.2.

表 1.2.2

月份	总数	男婴	女婴	生女频率
1	7 280	3 743	3 537	0.486
2	6 957	3 550	3 407	0.490
3	7 883	4 017	3 866	0.490
4	7 884	4 173	3 711	0.471
5	7 892	4 117	3 775	0.478
6	7 609	3 944	3 665	0.482