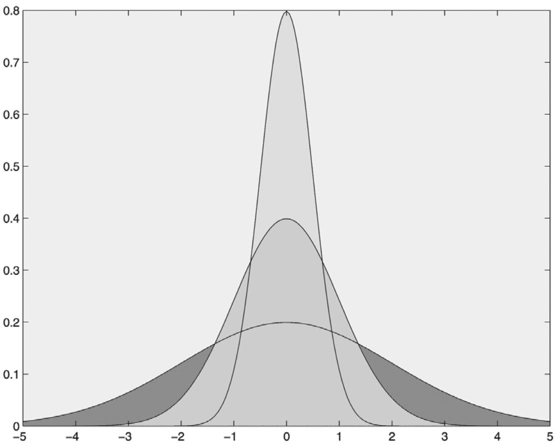
普通高中课程标准实验教科书

数学 第五册（必修）

教师教学用书 湘教版



湖南教育出版社

主　　编：张景中　黄楚芳
执行主编：李尚志
本册主编：邹捷中
编　　委：刘峰　康松林　李超贵

编 写 说 明

本书是《普通高中课程标准实验教科书·数学第五册（必修）》的教师教学用书．编写时按教科书分章、节安排，每章首先阐述了该章的教学目标、教材说明、教学建议、课时安排建议，然后按教科书分节编写，每节内容包括教材线索、教学目标、教材分析、教学建议、参考例题、相关链接，在每章的最后给出教科书中练习、习题和复习题的参考解答．

编写此书的目的是为了帮助教师更好地把握教科书，包括教材线索、教学目标、教材分析、内容结构及教学中应予以关注的重点和难点，所提教学建议及参考例题仅供教师在教学过程中参考．在相关链接中所提供的短文是编者精心撰写并与该章、节相关的内容，旨在扩大教师的知识视野，使教师用较高的观点把握教材，不要求学生掌握．

希望本书能成为教师使用教科书的好帮手，恳请广大教师在使用过程中提出宝贵意见和建议．谢谢！

编　者

目　录

第 11 章　算法初步

一、教学目标

1. 通过具体实例的分析，体会算法的思想，了解算法的含义，理解算法的一般特征，能用自然语言描述具体问题的算法；

2. 通过模仿、操作、探索，经历通过设计程序框图表达解决问题的算法过程，体验程序框图在解决问题中的作用，能理解并学习用程序框图描述具体问题的算法；

3. 在具体问题的解决过程中，理解程序框图的顺序结构、条件结构、循环结构等三种基本逻辑结构；

4. 结合具体问题，理解输入语句、输出语句、赋值语句、条件语句、循环语句等几种基本算法语句，理解它们与三种基本逻辑之间的关系；

5. 经历将具体问题的程序框图转化为算法语句，用伪代码描述算法的过程；

6. 通过中国古代及西方数学中具体的算法案例，进一步体会算法思想及算法的重要性和有效性，发展有条理的思考与表达的能力，提高逻辑思维能力，了解中国古代数学对世界数学发展的贡献.

二、教材说明

本章内容相对独立，但又和中学数学中的其他内容密切联系在一起，是高中数学课程中的新增内容，算法思想是非常重要的，但并不陌生和神秘. 随着现代信息技术飞速发展，算法在科学技术、社会发展中发挥着越来越大的作用，并日益融入社会生活的许多方面，算法思想已经成为现代人应具备的一种数学素养.

具体来说，本章的主要内容有：

（1）算法的概念：通过对解决具体问题过程与步骤的分析（如用“更相减损术”求两个正整数的最大公约数问题），体会算法的思想，了解算法的含义；

（2）算法程序框图的三种基本逻辑结构：顺序结构、条件结构、循环结构；

（3）基本的算法语句：输入语句、输出语句、赋值语句、条件语句、循环语句；

（4）典型的算法案例：用辗转相除法求最大的公约数，中国剩余定理，用二分法求方程的近似根，用秦九韶算法求多项式的值等.

本章内容的编排突出了以下特点：

1. 注重认知规律

本章内容强调通过案例引导学生认识算法的本质. 算法的概念并没有一个统一的定义，教材从丰富的实例出发，通过模仿、操作、探索，学习先用自然语言表达解决问题的过程，体会算法的基本思想，力求使学生能够对算法本质有所认识. 算法的表示按自然语言、程序框图、伪代码的顺序由易到难讲述了三种形式，这样便于学生理解和模仿. 教材通过简单的实例来说明程序框图和算法语言的使用，抓住了算法表示的核心内容，不追求完整. 算法案例的处理也遵循了这一原则，从算法的典型性、与以往知识的连续性和可接受性的角度出发，选择了中学生能够很容易理解的内容，从简单的例子到经典的问题，由易到难地讲述了四个典型算法案例，在阅读材料中还介绍了进位制，重在对案例的算法的分析，使学生能够通过案例的学习进一步理解算法的本质.

2. 突出横向联系

本章内容编写中还突出了与其他部分内容的联系，体现算法的基本思想. 比如一元二次方程求根过程、二元一次方程组的解法过程，用二分法求方程的近似解，数列求和等等. 力求通过这样的联系使学生认识到算法思想的重要性，并逐步能够应用算法思想解决一些实际问题.

3. 反映时代需求

算法内容反映了时代的特点，同时也是我国数学课程内容的新特色. 中国古代数学中蕴涵了丰富的算法思想，现代信息技术的发展使古老的算法焕发了前所未有的生机和活力，算法进入中学数学课程，反映了时代的要求，是中国古代数学思想在一个新的层次上的复兴.

三、课时安排建议

本章教学时间约需 12 课时，具体分配如下（仅供参考）

11.1 算法的概念　　1 课时

11.2 算法结构与程序框图

　　11.2.1 顺序结构　　1 课时

　　11.2.2 条件结构　　1 课时

　　11.2.3 循环结构　　1 课时

11.3 基本算法语句

　　11.3.1 输入、输出语句和赋值语句 1 课时

　　11.3.2 条件语句　　1 课时

　　11.3.3 循环语句　　1 课时

11.4 算法案例　　3 课时

小结与复习　　2 课时

四、教学建议

1. 抓住核心，准确把握教学要求

算法是解决“做什么”和“怎么做”的问题，一方面具有具体化、程序化、机械化的特点，同时又有高度抽象性、概括性和精确性. 对于一个具体算法而言，从算法分析到算法语言的实现，任何一个疏漏或错误都将导致算法的失败. 算法是思维的条理化、逻辑化. 算法思想可以贯穿于整个中学数学内容之中，有丰富的素材，而在算法的具体实现上又可以和信息技术相联系，因而，算法有利于培养学生理性精神和实践能力，是实施探究性学习的良好载体. 教学中应当把体会算法的基本思想、提高学生逻辑思维能力作为重点，应当以教材中提供的案例为载体，引导学生在设计程序框图、将程序框图转化为程序语句的实践中，体会算法的含义，学会如何用程序框图表达解决问题的思路，而不要拘泥于一种具体的语言的学习和程序设计.

2. 强调实践，尽量使用信息技术

算法的操作性很强，因此算法教学应当强调学生的动手实践. 教学中应当充分应用教材中提供的实例，使学生在解决具体问题的过程中学习一些基本逻辑结构和算法语句. 算法内容将数学中的算法与计算机技术建立了联系，为了有条理地、清晰地表达算法，往往需要将解决问题的过程整理成程序框图，为了能在计算机上实现，又要将自然语言或程序框图翻译成伪代码，稍加改造就可以在计算机平台上运行. 因此，如果能让学生上机，算法设计的整个过程就可以得到完整的体现，学生可以及时看到自己设计的算法的可行性、有效性，这不但可以很好地激发学生的兴趣，而且还能提高学习效果. 因此，有条件的学校，应鼓励学生尽可能上机尝试.

3. 横向联系，注重渗透算法思想

算法是高中数学课程中新增加的内容，其思想方法贯穿在所有的数学课程中，并不神秘和陌生，如解方程的算法、解不等式的算法、因式分解的算法等等都是我们熟知的内容. 所以在数学的学习过程中，应鼓励学生尽可能地运用算法解决相关问题，使程序化思想成为思考问题的习惯. 在教学中，要体现数学与算法的有机结合，在学习相应的内容（如制作随机数表、三角函数表、数列、不等式等）的过程中，有意识地引导学生体会算法思想，使他们看到数学在算法设计中的作用，以及掌握算法思想对于提高数学能力的重要性.

五、评价建议

在本章中，学生将在义务教育阶段初步感受算法思想的基础上，结合对具体数学实例的分析，体验程序框图在解决问题中的作用；通过模仿、操作、探索，学习设计程序框图表达解决问题的过程；体会算法的基本思想以及算法的重要性和有效性，发展有条理的思考与表达的能力，提高逻辑思维能力. 在教学中，要重视对学生数学学习过程的评价，包括学生参与数学活动的兴趣和态度、数学学习的自信、独立思考的习惯、合作交流的意识、数学认知的发展水平等方面. 要努力引导学生正确认识算法设计在数学以及计算机学科中的地位和作用，产生积极

的学习态度、动机和兴趣. 评价中应关注学生是否积极主动地参与算法设计活动，是否勤于思考，善于思考，逻辑思维是否严谨，以及是否愿意和能够与同伴交流算法设计的体会，与他人合作探究数学问题. 评价中应当重视考查学生能否利用算法的三种基本逻辑结构有条理地表达一般难度的算法设计问题的能力，不要拘泥于太复杂的算法设计过程，而要重视将实际生活中的问题抽象成计算机模型的抽象思维能力和表达能力.

11.1 算法的概念

教材线索

在以前的学习中，虽然没有出现算法这个名词，但实际上在数学教学中已经渗透了大量的算法思想. 本小节从学生熟知的四则运算的过程出发，指出完成这些工作都需要一系列程序化的步骤，也就是算法的思想，并指出算法的显著特点. 接着给出了两个例子，学会用自然语言表达算法步骤，层层深入地诠释算法的概念，揭示算法的特征.

教学目标

1. 通过具体实例，理解算法的概念与特点；

2. 学会用自然语言描述算法，体会算法思想，能说明解决简单问题的算法步骤；

3. 通过分析"更相减损术"的过程，学习有条理地、清晰地表达解决问题的步骤，培养学生逻辑思维能力与表达能力，发展从具体问题中提炼算法思想的能力.

教材分析

1. 重点

(1) 算法的概念；

(2) 用自然语言描述算法；

(3) 算法的基本特征.

2. 难点

(1) 如何用自然语言来描述算法；

(2) 对算法基本特征中有效性和有穷性的理解；

(3) 对"更相减损术"原理的理解.

3. 对算法概念的理解

算法不仅是数学及其应用的重要组成部分，也是计算机科学的重要基础. 我们在解决某些问题时，需要设计出一系列可操作或可计算的步骤，通过实施这些步骤来解决问题，通常把这些步骤称为解决这些问题的算法. 也就是说，算法指可以用计算机解决的某一类问题的程序或步骤，这些程序或步骤必须是明确和有效的，而且能够在有限步之内完成.

算法有以下几个主要特征：

①确定性：算法中的每一步操作的内容和顺序必须含义确切，不能有二义性；

②有效性：算法中的每一步操作都必须是可执行的，也就是说算法中的每一步都能通过手

工和机器在有限时间内完成，这称之为有效性；

③有穷性：算法中执行的步骤总是有限次数的，不能无休止地执行下去.

有些人提出算法还应具有“通用性”，所谓的“通用性”即一个算法应该适用于求某一类问题的解，而不是用来解决一个具体问题. 对于初学者来说，算法的通用性不宜过分强调，以避免在一开始就要把解决的问题复杂化. 在解决具体问题的过程中，可以先针对具体问题设计算法，在设计的过程中或在设计完成之后，再考虑如何修改或推广该算法以解决类似的一类问题，以达到通用性的要求.

在这几个特征中，有效性和有穷性是算法的最重要的两个特征.

算法的有效性指的是算法中的每一步运算和操作必须是相当基本的，从而是能实现的. 但是，所谓的基本和能实现都是相对而言的. 例如，对于一元二次方程求根的操作，在解决其他问题时，可以直接作为一个步骤用在算法中. 但如果要求写出一元二次方程求根的算法时，我们就不认为它是基本的，而是要写出求根的各个步骤.

算法的有穷性是十分重要的，是成功完成任务的必要条件. 特别在算法的循环结构中，一定要考虑好何时终止循环，转入下一步操作. 由于算法的循环结构在课程中出现较晚，学生在刚开始时不易体会算法有穷性的重要，这关系不大．但在课程进行到适当的时候，应不失时机地提醒学生注意算法有穷性的要求.

教学建议

算法是中学数学中的一个新的概念，它没有精确化的定义，教材中只是作了如下描述：“算法通常是指由有限多个步骤组成的求解某一类问题的通用方法.”所以，教学中，只要使学生明确算法实际上就是解决某一个或某一类问题的程序化方法. 中学数学中的算法强调具体算法所蕴含的算法思想，重点在于培养学生的算法意识，这是算法教学中始终应该注意的.

在教学过程中不要求讲解算法的严格定义，也不要求学生去记忆算法的特征，但教师应注意从算法的几个特征去解释算法的含义，并引导学生从这些方面去体会算法的含义. 教材上的例题是学生在小学就接触过的，学生并不陌生，选择这样的问题，一方面是希望打破学生对算法的神秘感，另一方面是希望把重点放在对算法的理解上，而不是算法所涉及的问题本身. 建议增加讲解不同类型的例题，以加深学生对算法的理解，并进一步学习如何用自然语言来描述算法.

教材中的两个例子都与约数有关，其中例 1 已讲述得较为清晰，这里不再赘述. 旁注中的问题“求正整数 n（$n>2$）的大于 1 的最小约数”的算法步骤可以表述为：

S1：设 $k=2$；

S2：求出 n 除以 k 的余数 r，若 $r=0$，则执行 S4；否则执行 S3；

S3：k 的值增加 1，返回 S2；

S4：k 就是所求的数.

而例 3 中的“更相减损术”的原理较难理解.“更相减损术”是我国古代数学中求两整数最大公约数的方法. 古代数学专著《九章算术》卷一在谈到分数分子分母约去公因数有“可半者半之，不可半者，置分母子之数，以少减多，更相减损求其等也，以等数约之.”翻译成现代汉语语言为：(1) 任意给出两个正数，判断它们是否都是偶数，若是，用 2 约简；若不是，执行第二步；(2) 以较大的数减去较小的数，接着把较小的数与所得的差比较，并以大数减小数. 继续这个操作，直到所得的数与较小的数相等为止，则这个数（等数）就是所求的最大公约数. 这也是教材旁注中的口诀“大数减小数，用差替大数；依此反复做，直到相等数.”的具体含义.

由于学生是初次接触算法，在进行例题教学时，只要求学生能认识算法步骤，会用自然语言写简单具体问题的算法步骤，不要求在算法的“优化”上做文章.

例题分析

为了更好的说明算法的概念，可以继续补充下面的例子加深学生对算法的理解.

例 1. 判断一个大于 1 的整数 n（比如 18 441 803）是否为质数（5 347×3 449=18 441 803)?

算法分析：

S1：判断 n 是否等于 2，若 $n=2$，则 n 是质数；若 n 大于 2，则执行 S2；

S2：依次从 2～$(n-1)$ 中检验是不是有 n 的因数，即整除 n 的数，如果有这样的数，则 n 不是质数，否则 n 是质数.

说明：根据算法的定义，上述过程是明确的、确定的、而且能在有限步内完成，它就是一个算法. 尽管 n（比如 18 441 803）可能是一个很大的数，人工很难在短时间内完成，但是借助于机器可以在较短时间完成（也许只有几毫秒).

尽管上述算法可以解决我们的问题，但它并不是最佳算法.

我们可以这样改进：

S1：判断 n 是否等于 2，若 $n=2$，则 n 是质数；若 n 大于 2，则执行 S2；

S2：设 $k=2$；

S3：求 n 除以 k 的余数 r；

S4：k 的值增加 1；

S5：判断 $k \geqslant n$ 或 $r=0$ 是否成立，若成立，执行 S6；若不成立，返回 S3；

S6：判断 $r=0$ 是否成立，若成立，n 不是质数，否则 n 是质数.

我们看改进的算法，若一个数不是质数，则不必判断到 $n-1$ 即可得到结果. 例如 18 441 803判断到 3 449 即知它是一个合数.

事实上，我们还可以进一步优化算法，在 S5 中，只须判断 $k \geqslant \frac{n}{2}$ 或 $r=0$ 是否成立即可，甚至只须判断 k 是否大于等于 $\sqrt{n}$ 或 $r=0$ 即可．这里牵扯到算法复杂度问题，不必向学生介绍．

这个例题说明，解决一类问题的算法不是唯一的．

例 2． 求一个整数序列的最大值．

初看起来是很简单的事，如从 8 个，甚至 10 个、20 个整数中找最大数，可能容易发现它，但若是 100 个，甚至 1 000 个，10 000 个呢？一般地，可以设第一个是最大值，然后和第二个比较，再取出最大值和第三个比较，一直进行下去，直到最后一个数．

算法分析：

S1：假定这个序列的最大值为 m，m 赋值为序列中的第一个数；

S2：将序列中的下一个整数与 m 比较，如果它大于 m，那么就用这个数赋值给 m；

S3：如果序列中还有其他整数，重复 S2；

S4：一直到序列中没有可比数为止，这时的 m 就是这个序列的最大值．

相关链接

算法一词的由来

Algorithm（算法）一词本身就十分有趣．初看起来，这个词好像是某人打算要写“Logarithm”（对数）一词但却把头四个字母写的前后颠倒了．这个词一直到 1957 年之前在 Webster's New World Dictionary（《韦氏新世界词典》）中还未出现，我们只能找到带有它的古代涵义的较老形式的“Algorism”（算术），指的是用阿拉伯数字进行算术运算的过程．在中世纪时，珠算家用算盘进行计算，而算术家用算术进行计算．中世纪之后，对这个词的起源已经拿不准了，早期的语言学家试图推断它的来历，认为它是从把 algiros（费力的）＋arithmos（数字）组合起来派生而成的，但另一些人则不同意这种说法，认为这个词是从“喀斯迪尔国王 Algor”派生而来的．最后，数学史学家发现了 algorism（算术）一词的真实起源：它来源于著名的 PersianTextbook（《波斯教科书》）的作者的名字 Abu Ja'far Mohammed ibn Mûsâal－Khowârizm（约公元前 825 年）——从字面上看，这个名字的意思是“Ja'far 的父亲，Mohammed 和 Mûsâ 的儿子，Khowârizm 的本地人”．Khowârizm 是前苏联 XИBA（基发）的小城镇．Al－Khowârizm 写了著名的书 Kitab al jabr w'al－muqabala《复原和化简的规则》）；另一个词，“algebra”（代数），是从他的书的标题引出来的，尽管这本书实际上根本不是讲代数的．

逐渐地，“algorism”的形式和意义就变得面目全非了．如牛津英语字典所说明的，这个词是由于同 arithmetic（算术）相混淆而形成的错拼词，由 algorism 又变成 algorithm．一本早期

的德文数学词典 Vollstandiges Mathematisches Lexicon（《数学大全辞典》），给出了 Algorithmus（算法）一词的如下定义：“在这个名称之下，组合了四种类型的算术计算的概念，即加法、乘法、减法、除法”. 拉丁短语 algorithmus infinitesimalis（无限小方法），在当时就用来表示 Leibniz（莱布尼兹）所发明的以无限小量进行计算的微积分方法.

1950 年左右，algorithm 一词经常地同欧几里得算法（Euclid's algorithm）联系在一起，这个算法就是在欧几里得的《几何原本》（Euclid's Elements，第Ⅶ卷，命题ⅰ和ⅱ）中所阐述的求两个数的最大公约数的过程（即辗转相除法）.

11.2 算法结构与程序框图

11.2.1 顺序结构

教材线索

教材先从上节中例题的算法表示入手，说明算法步骤有明确的顺序性. 用自然语言表示算法时，特别是需要表达那些在一定条件下才会被执行的步骤，以及一些在一定条件下会被重复执行的步骤时，会显得困难，而且不直观、不明确，从而过渡到介绍用程序框图表示算法. 教材在介绍了常用的程序框图基本元件后，再逐一安排对三种基本逻辑结构的学习. 本节通过两个典型例题介绍最基本、最简单的顺序结构.

教学目标

1. 理解程序框图的概念，能读懂用程序框图描述的算法，初步理解用程序框图表示算法的优越性；

2. 掌握常见程序框图基本元件及其功能；

3. 会用程序框图表示简单顺序结构的算法；

4. 经历由自然语言表示算法转化为用程序框图表示算法的过程，体会顺序结构；

5. 通过实例，引导学生由问题求出算法，培养学生的逻辑思维能力.

教材分析

1. 重点

(1) 程序框图基本元件及其功能，算法的顺序结构的概念；

(2) 画顺序结构的程序框图；

(3) 算法顺序结构的判别和应用.

2. 难点

(1) 算法顺序结构的应用；

(2) 理解“鸡兔同笼”问题的算法原理.

3. 顺序结构分析

顺序结构是最简单的算法结构，是任何一个算法中必不可少的结构，它表示语句和语句之

间，程序框与程序框之间是按从上到下的流线顺序依次进行的.

4.“鸡兔同笼”问题的理解

“鸡兔同笼”是一类有名的中国古算题，最早出现在《孙子算经》中. 许多算术应用题都可以转化成这类问题，因此很有必要学会它的解法和思路.

对于这个问题，孙子提出了大胆的设想：他假设砍去每只鸡、每只兔一半的脚，则每只鸡就变成了“独脚鸡”，而每只兔就变成了“双脚兔”. 这样，“独脚鸡”和“双脚兔”的脚就由 94 只变成了 47 只；而每只“鸡”的头数与脚数之比变为 1 比 1，每只“兔”的头数与脚数之比变为 1 比 2. 由此可知，有一只“双脚兔”，脚的数量就会比头的数量多 1. 所以，“独脚鸡”和“双脚兔”的脚的数量与他们的头的数量之差，就是兔子的只数，即：47－35＝12（只）；鸡的数量就是：35－12＝23（只）.

这一思路新颖而奇特，其“砍足法”也令古今中外数学家赞叹不已. 这种思维方法叫化归法，就是在解决问题时，先不对问题采取直接的分析，而是将题中的条件或问题进行变形，使之转化，直到最终把它归结成某个已经解决的问题.

教学建议

教材中首先将上节两个例题的算法用程序框图呈现出来，目的是让学生在比较中感受用程序框图表示算法的优越性，教学时应引导学生对照算法的自然语言表述，学会阅读这两个程序框图，理解不同形状的程序框代表不同的含义.

程序框图是算法的一种表现形式，算法可以用自然语言表示，也可以程序框图描述. 通常是先写出算法的算法步骤，然后再转化为对应的程序框图. 终端框、输入输出框、处理框、判断框及流程线等是是组成程序框图的基本元件，它们有各自的含义. 教学时应使学生规范地使用这些图形符号，不得改变它们的意义或者随意增加其他图形. 一个完整的程序框图一定要以终端框开始，同时又以终端框结束. 流程线是带有方向箭头的线，用来连接程序框，表示算法的流程. 在教学时，要注意判断框是唯一有两条流出线的程序框.

顺序结构理解较为简单，主要突出它的作用. 另外，在讲解顺序结构中要注意讲解算法的思想，培养学生的思维能力，为后续学习作铺垫. 教材中的例 1 选择用线段中点坐标公式求线段的中点坐标，问题本身很简单，目的在于理解算法的顺序结构，学会画程序框图来描述算法.

例 2 是古老的“鸡兔同笼”问题，教材选用方程组解法，继而在旁注中提出一般的算法，目的在于让学生明白算法其实是贯穿在数学各领域中的. 在教学中，引导学生回答旁注中的问题“想一想，为什么可以这样做?”时，除可以向学生介绍前面提到的孙子的“砍足法”外，还可以形象地分析：饲养员一声令下，所有的鸡都抬起左脚，扮演金鸡独立，所有的兔子都抬起前脚，直立起来，此时有 47 只脚着地，如果把兔头也看作鸡头，就只会有 35 只脚着地，多出的 12 只脚哪来的呢？当然是 12 只兔子的!

学生初学画程序框图，可能会感到比较难于操作，应该给予适当的指导. 画程序框图应该

遵循一些共同的规则：

（1）使用标准的程序框符号；

（2）按照从上到下、从左到右的方向画；

（3）除判断框外，大多数框图符号只有一个进入点和一个退出点.

（4）程序框内的文字描述语言要尽可能简练、清楚.

用程序框图表示算法，直观、形象，容易理解，通常说“一图胜万言”，就是说用程序框图能更清楚地展现算法的逻辑结构.

本节最后解一般的二元一次方程组$\begin{cases} a_1x+b_1y=c_1, \\ a_2x+b_2y=c_2 \end{cases}$（$a_1b_2-a_2b_1\neq 0$）的算法如下：

S1：输入方程组系数 a_1，b_1，c_1，a_2，b_2，c_2；

S2：计算 $x=\dfrac{b_2c_1-b_1c_2}{a_1b_2-a_2b_1}$，$y=\dfrac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}$；

S3　输出计算结果 x、y.

例题分析

例交换两个变量 A 和 B 的值，并输出交换后的值.

算法步骤如下：

S1：输入 A，B 的值；

S2：把 A 的值赋给 x；

S3：把 B 的值赋给 A；

S4：把 x 的值赋给 B；

S5：输出 A，B 的值.

程序框图如图 11－1 所示.

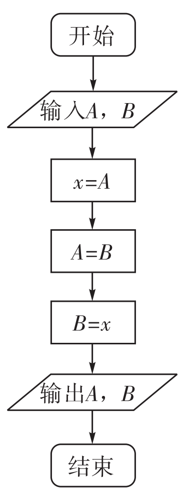


图 11－1

相关链接

算法的表示方法

所谓“算法”，粗略地讲，是为解决一个特定问题而采取的确定的有限的步骤. 广义来说，做任何事情都有其算法，例如，一本歌谱可以说就是歌曲的算法，因为它规定了歌唱者应如何唱（先唱什么，后唱什么，什么音阶什么音符），一张太极拳打法图解也是一个“太极拳算法”. 凡做任何事都需要先确定算法，然后去实现这个算法以达到目的. 当然本书只涉及计算机算法，即计算机能执行的算法. 例如指定让计算机“去买一包烟”，计算机是无法实现的（至少目前如此），而让计算机求一个一元二次方程的解是可以实现的. 算法的表示方法可以有以下几种：

1. 用自然语言表示

可以用中文或英文句子来说明该做什么. 但是用自然语言表示一般比较冗长，而且有“歧义性”（即对同一段文字，不同的人会有不同的理解），如：“张三对李四说他的儿子考上了大学”，究竟指的是张三的儿子还是李四的儿子呢？这段文字中难以确定.

2. 用一般流程图表示

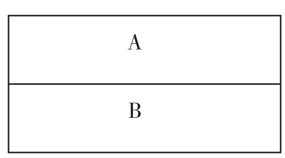
起止框表示算法的开始和结束；处理框，主要用来表示“赋值”等操作；判断框用来根据给定的条件是否满足决定执行二条路径中的一条路径；输入输出框用来表示输入输出操作；流程线的箭头表示流程的方向，这种流程图表示算法的好处是：用图形来表示算法流程，直观形象，各种操作一目了然，而且不会产生“歧义性”，流程清晰. 缺点是占用面积大，而且由于允许使用流程线，使流程任意转移，容易使人弄不清算法流程的意思.

3. 用N－S结构化流程图表示

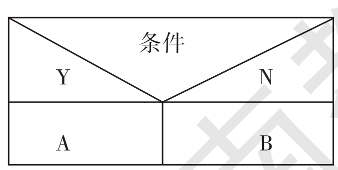
N－S流程图的主要特点是取消了流程线，即不允许流程任意转移，而只能从上到下顺序进行. 它规定了几种基本结构作为构造算法的基本单元. N－S图是美国学者 I. Nassi 和 B. Schneiderman在1973年提出的方法的基础上完成的一种基于结构化的流程图，因此称为N－S流程图.

N－S流程图基本结构如下：

（1）顺序结构

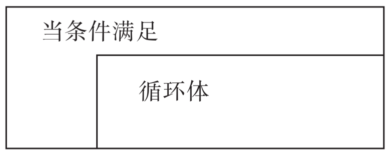


（2）选择结构

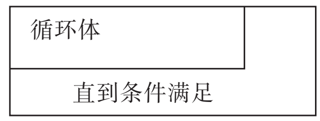


（3）循环结构

当型循环



直到循环



4. 用伪代码表示算法

所谓伪代码是指用介于自然语言和计算机语言之间的一种代码来描述算法的方法. 它的表示形式比较灵活自由，而且由于与计算机语言比较接近，因此可以比较容易地转换成计算机程序. 显然汉字不如英文方便，因为用英文单词很容易换成高级语言程序. 用伪代码表示算法可以不必画流程图，比较省时省力，而且在修改算法时比修改流程图方便、但初学者一般喜欢使用流程图表示算法，因为它比较形象直观、能清晰地表示出流程，即使出现错误也容易发现.

11. 2. 2　条件结构

教材线索

本小节开门见山直接介绍了条件结构的概念和在程序框图中的基本表现形式，然后由浅入深、循序渐进的通过三道例题介绍了条件结构的应用，让学生初步感受条件结构在算法中的关键作用.

教学目标

1. 通过实例介绍，认识条件结构，能读懂包含条件结构的算法程序框图；

2. 掌握条件结构的应用，学会绘制包含条件结构算法的程序框图；

3. 能将具体问题抽象出含有条件结构的算法，并用程序框图描述，培养学生抽象概括能力、语言表达能力及思维的严密性.

教材分析

1. 重点

（1）算法的条件结构的概念和程序框图；

（2）算法条件结构的判别、作用和应用.

2. 难点

（1）算法条件结构的作用和应用；

（2）条件结构的嵌套.

3. 条件结构的理解

条件结构又称分支结构，是指在算法中通过对条件的判断，根据条件是否成立而选择不同流向的算法结构. 在许多算法中，需要对问题的条件作出逻辑判断，判断后依据条件是否成立而进行不同的处理方式，这就需要用条件结构来实现算法. 分类是算法中经常发生的事情，条件结构的主要作用就是实现分类.

教材中介绍的条件结构主要有两种，一种是在两个分支中都包含算法的步骤，符合条件就执行步骤 A，否则执行步骤 B；另一种是在一个分支中包含算法步骤 A，另一个分支中不包含算法步骤，符合条件就执行步骤 A，否则就执行条件结构后面的步骤. 在分类不止两类的情况下，通常要用到在条件结构中再嵌入条件结构的方法来实现，称为条件结构的嵌套或称复合条件结构.

以前学生学过的很多算法问题都可以用条件结构来表示，比如方程求根、分段函数等，可以在这方面多给学生思考的空间.

教学建议

在刚开始学习时，要引导学生通过模仿、操作、探索，经历通过设计程序框图表达解决问题的过程，特别在本节条件结构的教学中，要尽量让学生自己挖掘出所需的判断条件，并清楚流程图的“行进路线”.

分段函数的求值问题，往往要用到条件结构实现算法，这一问题是学生最熟悉的. 在教学时，可以把例 1 和例 2 定位在阅读层面，把例 3 定位在操作层面，这也是教材编写者的意图. 先通过实例认识条件结构，并结合旁注中的问题，让学生模仿例题的程序框图，尝试改变判断框中的条件，或者改变嵌套条件的位置，画出不同的程序框图.

教材中例 1 和例 3 的程序框图所包含的条件结构都是第一种类型，例 2 的程序框图是条件结构的嵌套. 其实例 1 的程序框图稍加改造就可以变成另一种条件结构的形式（如图 11－2）：

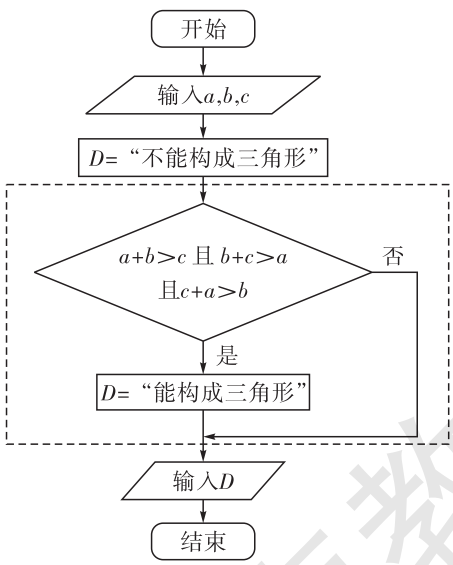


图 11－2

另外，例 1 中对已知三边长度来判定这三边是否能构成三角形还有其他判定方法，譬如利用“三角形的两边之差小于第三边”，很明显这两种方法是等效的.

例 2 中两个判断框的条件也是可以改变的，但算法流程的走向是与条件相关联的. 教学中，应让学生明白可以在“$\Delta>0$”、“$\Delta=0$”、“$\Delta<0$”三个条件中任意选择两个作为判断条件都是可行的.

例 3 的教学中可以选择变式问题：

已知函数 $f(x)=\begin{cases} x^2-1, (x>0), \\ 0, (x=0), \\ 2x-1, (x<0), \end{cases}$ 设计一个算法，输入 x 的值，输出 y 的值.

引导学生画出程序框图，并进一步分析复合条件结构中的内、外层嵌套关系（如图 11－3 所示）. 进一步，还可引导学生作如下思考：当 x 取 2，0，－2 时，算法流程分别沿着哪条路线运行？以加深学生们对复合条件结构的理解.

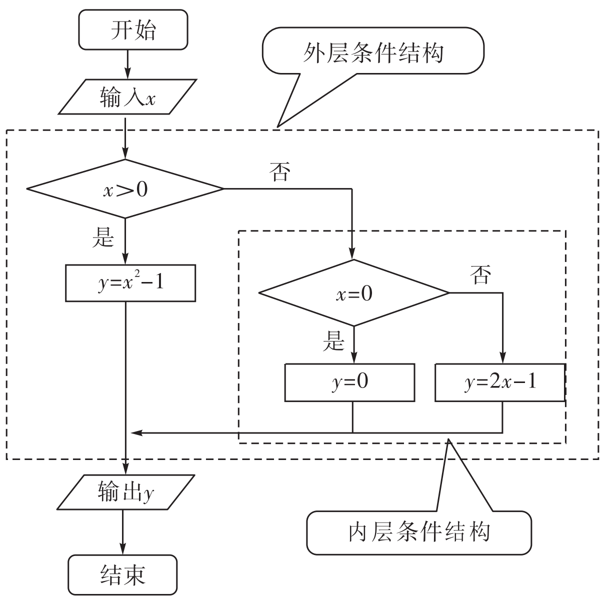


图 11－3

例题分析

作为教材例 1 的深入研究，补充下面的问题.

例. 已知三角形三边的长分别为 a，b，c（可以构成三角形），判定此三角形是钝角三角形、锐角三角形还是直角三角形，请给出算法.

算法分析：

一种直接的办法就是用余弦定理来判定三角形的三个内角，若有某个角的余弦值为负数，则此三角形为钝角三角形；若有某个角的余弦值为零，则此三角形为直角三角形；若三个角的余弦值均为正数，则此三角形为直角三角形.

显然，上述办法是比较复杂的，那么有没有较简便的办法呢？

我们知道，判断一个三角形是钝角三角形、锐角三角形还是直角三角形，主要是看这个三角形的最大内角，同时我们由正弦定理也知道，在三角形中大边对应大角（即长度最长的边对应的内角最大）. 因此，我们可以先比较三角形三边的大小，然后再对最长的边所对应的角用余弦定理来判定. 我们还可以用勾股定理来进行判断，这种方法更简便.

用勾股定理来进行判定的算法步骤如下：

S1：输入三角形三边 a，b，c；

S2：比较三角形三边的大小，将最长边用 a 表示；

S3：若 $a^2>b^2+c^2$，则输出“$\triangle ABC$ 为钝角三角形”；若 $a^2=b^2+c^2$，则输出“$\triangle ABC$ 为直角三角形”；否则输出“$\triangle ABC$ 为锐角三角形”，

程序框图如图 11－4 所示.

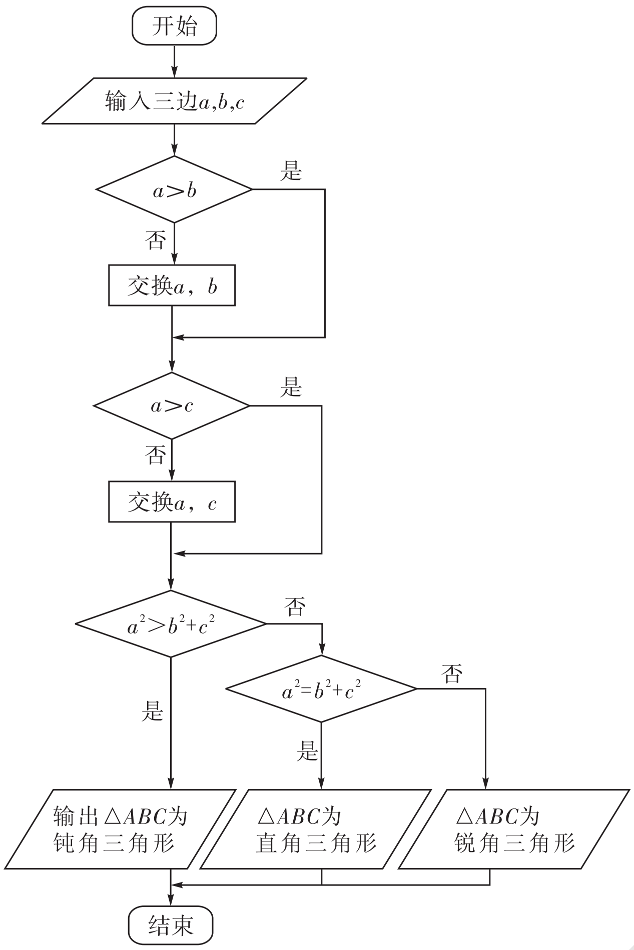


图 11－4

在本程序框图中，第四个框的“交换 a，b”也可用“$t=a$，$a=b$，$b=t$”实现，在上节补充例题中已经讲到. 第六个框同理，这样经过处理后，保证了 a 即为最长的边.

要注意的是，本程序框图中前面两个条件“$a>b$”和“$a>c$”是并列的，没有形成嵌套关系，也没有先后顺序，可对任一条件先进行判断，对结果没有影响. 而后面的两个条件“$a^2>b^2+c^2$”和“$a^2=b^2+c^2$”就形成了嵌套关系.

相关链接

算法的复杂性

算法的复杂性是算法效率的度量，在评价算法性能时，复杂性是一个重要的依据. 算法的复杂性的程度与运行该算法所需要的计算机资源的多少有关，所需要的资源越多，表明该算法的复杂性越高；所需要的资源越少，表明该算法的复杂性越低.

计算机的资源，最重要的是运算所需的时间和存储程序和数据所需的空间资源，算法的复杂性有时间复杂性和空间复杂性之分.

算法在计算机上执行运算，需要一定的存储空间存放描述算法的程序和算法所需的数据，计算机完成运算任务需要一定的时间. 根据不同的算法写出的程序放在计算机上运算时，所需要的时间和空间是不同的. 算法的复杂性是对算法运算所需时间和空间的一种度量. 不同的计算机其运算速度相差很大，在衡量一个算法的复杂性要注意到这一点.

对于任意给定的问题，设计出复杂性尽可能低的算法是在设计算法时考虑的一个重要目标. 另外，当给定的问题已有多种算法时，选择其中复杂性最低者，是在选用算法时应遵循的一个重要准则. 因此，算法的复杂性分析对算法的设计或选用有着重要的指导意义和实用价值.

在讨论算法的复杂性时，有两个问题要弄清楚：

(1) 一个算法的复杂性用怎样的一个量来表达；

(2) 怎样计算一个给定算法的复杂性.

找到求解一个问题的算法后，接着就是该算法的实现. 至于是否可以找到实现的方法，取决于算法的可行性和计算的复杂性，该问题是否存在求解算法，能否提供算法所需要的时间资源和空间资源.

11.2.3 循环结构

教材线索

本节先介绍循环结构的有关概念，阐述两种循环结构类型的区别与联系，以及在程序框图中的基本表现形式，然后通过两个例子分别介绍当型循环和直到型循环的应用.

教学目标

1. 认识循环结构，能读懂含有循环结构的程序框图；

2. 通过实例介绍，模仿设计含有循环结构的程序框图的过程，熟悉循环结构的应用；

3. 能将具体问题抽象出含有循环结构的算法，并用程序框图描述，培养学生抽象概括能力、语言表达能力及逻辑思维能力.

教材分析

1. 重点

（1）循环结构的概念和包含循环结构的程序框图的理解；

（2）循环结构的判别、作用和应用；

（3）条件结构与循环结构的区别与联系.

2. 难点

（1）循环体的确定，计数变量与累加变量的理解；

（2）两种循环结构的程序框图的画法.

3. 对循环结构的理解

计算机的优势就是能够以极快的速度进行重复计算，循环结构的重要性也在于此. 因此，在设计算法时，要充分利用循环结构所提供的方便.

在一些算法中，也经常会出现从某处开始，按照一定条件，反复执行某一处理步骤的情况，这种结构称为循环结构. 反复执行的处理步骤称为循环体.

在循环结构中，通常都有一个起到循环计数作用的变量，称为计数变量，这个变量的取值一般都含在循环条件中，用来控制执行或终止循环体.

当型循环是在每次执行循环体前对控制循环条件进行判断，当条件满足时执行循环体，不满足则停止.

直到型循环是在执行了一次循环体之后，对控制循环体进行判断，当条件不满足时执行循环体，满足则停止.

当型循环与直到型循环的区别：

①当型循环可以不执行循环体，直到型循环至少执行一次循环体；

②当型循环先判断后执行，直到型循环先执行后判断；

③对同一算法来说，当型循环和直到型循环的条件互为否定条件.

循环结构主要用在反复做某项工作的问题中，循环结构本质上是判断和处理的结合，可以是先判断，再处理，即当型循环；也可以是先处理，再判断，即直到型循环. 对同一算法来说，用当型循环和直到型循环描述都是可以的.

画循环结构程序框图前：①确定循环变量和初始条件；②确定算法中反复执行的部分，即循环体；③确定循环的转向位置；④确定循环的终止条件.

条件结构与循环结构的区别与联系：

区别：条件结构通过判断分支，只执行一次；循环结构通过条件判断可反复执行.

联系：循环结构是通过条件结构来实现的.

教学建议

循环结构是算法基本逻辑结构中的重点，也是难点．应用循环结构来实现反复执行的计算是一种新的思想和方法，刚开始时不容易掌握，学习时有一定的困难，建议结合具体实例来进行讲解，主要是要求学生习惯这种新的思维方式.

在例1的教学中，应注意讲清楚以下几点：

（1）算法流程按箭头所指方向进行，在判断框处的条件是“$k\leqslant 100$”，如果 $k>100$，则按“是”所指方向进行循环，否则按“否”所指方向进行．当然，我们也可以设定条件“$k>100$”，如果 $k\leqslant 100$，则按“否”所指方向进行循环，否则按“是”所指方向进行．在程序框图中，交换“是”、“否”位置就可以了，二者是等价的.

（2）k 叫作计数变量，用于记录循环次数，同时它的取值还可以用来判断循环是否中止．循环第一圈时 k 的值为1，第二圈时 k 的值为2，……，第100圈时 k 的值为100．这种变化是通过 $k=k+1$ 来实现的，其含义是“后来的 k 值＝原 k 值＋1”．每循环一圈，$k=k+1$ 被执行一遍，k 的值就增加了1.

（3）S 叫作累加变量，用于记录累加结果．循环第1圈在 S 上加1，循环第2圈在 S 上加 $\frac{1}{2}$，循环第3圈在 S 上加 $\frac{1}{3}$；……，循环第 k 圈在 S 上加 $\frac{1}{k}$，……，循环第100圈在 S 上加 $\frac{1}{100}$.

“循环第 k 圈在 S 上加 $\frac{1}{k}$” 可以用 $S=S+\frac{1}{k}$ 表示，把这时 $\frac{1}{k}$ 的值加到 S 上．当循环到100圈时，S 的值就是 $1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{100}$ 的和.

为了使学生能够真正理解循环结构，可以要求他们在下表中填入相应的数字.

	累加变量					计数变量			
	S	=	S	+	$1/k$	k	=	k	+ 1
第1圈		=		+			=		+ 1
第2圈		=		+			=		+ 1
第3圈		=		+			=		+ 1
第4圈		=		+			=		+ 1
第5圈		=		+			=		+ 1
第6圈		=		+			=		+ 1
………									
第100圈		=		+			=		+ 1

教材中例 1 和例 2 分别使用了循环结构的两种类型，例 1 属于当型循环，而例 2 属于直到型循环. 在教学中可以要求学生改用另一种类型循环结构来表达，体会两种循环结构类型之间的相互转换.

下面是分别用当型循环和直到型循环两种循环结构描述的同一个具体实例的程序框图，一个饿汉吃饼的情况，如果饥饿的话就吃一张饼，直到吃饱为止.

当型循环

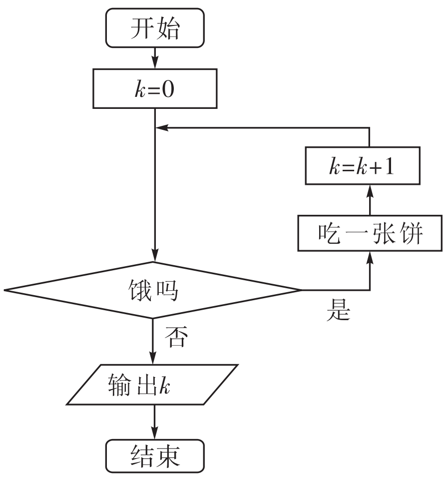


图 11－5

直到型循环

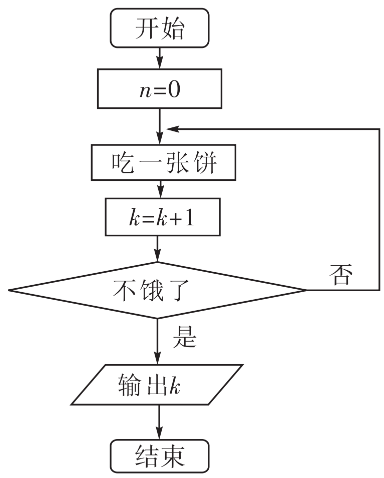


图 11－6

总之：当型循环结构的特点是“当满足条件时就循环”；直到型循环结构的特点是“直到满足条件时退出”.

例题分析

教材例 2 变式问题：

例. 某人购买了一辆价值 25 万元的轿车，汽车将以每年 20％的速度折旧，请用算法描述购车 1 到 5 年中汽车的价值变化，并用程序框图表示出来.

算法分析：按照题意，设汽车当年价值为 P，汽车每年折旧 20％，说明一年后汽车价值是上一年的 80％，即 $0.8P$，反映汽车价值变化即是这样一个循环过程，可以用循环结构来实现.

算法步骤如下：

S1：初始化变量 $P=25$，$k=1$；

S2：计算并输出下一年汽车价值 P，k 的值增加 1；

S3：判断 k 是否大于 5，若是，结束算法；否则返回 S2.

程序框图如图 11－7 所示.

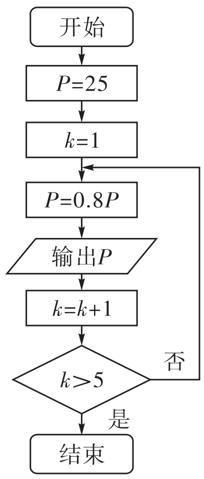


图 11－7

相关链接

刘　徽

刘徽（生于公元 250 年左右），是中国数学史上一个非常伟大的数学家，在世界数学史上，也占有重要的地位. 他的杰作《九章算术注》和《海岛算经》，是我国最宝贵的数学遗产.

《九章算术》约成书于东汉之初，共有 246 个问题的解法. 在许多方面：如解联立方程，分数四则运算，正负数运算，几何图形的体积面积计算等，都属于世界先进之列，但因解法比较原始，缺乏必要的证明，而刘徽则对此均作了补充证明. 在这些证明中，显示了他在多方面的创造性的贡献. 他是世界上最早提出十进小数概念的人，并用十进小数来表示无理数的立方根. 在代数方面，他正确地提出了正负数的概念及其加减运算的法则；改进了线性方程组的解法. 在几何方面，提出了“割圆术”，即将圆周用内接或外切正多边形穷竭的一种求圆面积和圆周长的方法. 他利用割圆术科学地求出了圆周率 $\pi=3.14$ 的结果. 刘徽在割圆术中提出的“割之弥细，所失弥少，割之又割以至于不可割，则与圆合体而无所失矣”，这可视为中国古代极限观念

的佳作.

《海岛算经》一书中，刘徽精心选编了九个测量问题，这些题目的创造性、复杂性和富有代表性，都在当时为西方所瞩目.

刘徽思想敏捷，方法灵活，既提倡推理又主张直观．他是我国最早明确主张用逻辑推理的方式来论证数学命题的人.

11.3 基本算法语句

11.3.1 输入、输出语句和赋值语句

教材线索

教材在安排学习了算法的两种描述方法后指出，要使算法成为计算机解决问题的工具，还必须学习用算法语句来描述算法，从而引入了容易改写成各种计算机程序的算法语言，即伪代码. 教材对照三种基本逻辑结构，分三节介绍基本算法语句，本小节采用先阅读理解，后讲述规范，再操作实践的方法学习输入、输出语句和赋值语句.

教学目标

1. 理解输入语句、输出语句和赋值语句的功能和一般格式；

2. 理解变量的概念，掌握变量的赋值；

3. 通过实例，初步了解并掌握将算法的描述变成伪代码的过程，比较自然语言、程序框图和伪代码表示算法的区别和联系；

4. 进一步体会算法的基本思想，能准确地运用输入语句、输出语句和赋值语句.

教材分析

1. 重点

（1）输入语句、输出语句和赋值语句的功能和一般格式；

（2）将算法的描述变成伪代码的过程，伪代码的书写.

2. 难点

（1）赋值语句的理解；

（2）伪代码的书写.

3. 关于伪代码

伪代码（Pseudocode）是一种算法描述语言，它不是一种现实存在的编程语言. 使用伪代码的目的是为了使被描述的算法可以容易地以任何一种编程语言（Pascal，C，Java，etc 等）实现. 它可能综合使用多种编程语言中语法、保留字，甚至会用到自然语言. 因此，伪代码必须结构清晰，代码简单，可读性好，并且类似自然语言. 计算机科学在教学中通常使用伪代码，以使得所有的程序员都能理解.

在伪代码书写上，用缩进表示程序的块结构，可以大大提高代码的清晰性. 同一模块的语句有相同的缩进量，次一级模块的语句相对于其父级模块的语句缩进. 简单来说，伪代码相当于为了实现自己算法思想所写的一种能令人容易明白的接近源代码的算法表现符号.

在伪代码中的运算符或函数，和我们平常用的有些有所不同，常见的有：

符号	作用及示例	符号	作用及示例
+	加法运算，与通常写法相同	−	减法运算，与通常写法相同
*	乘法运算，如 2*3=2×3=6	/	除法运算，如 6/2=6÷2=3
∧	乘方运算，如 6∧2=6^2=36	MOD	求余运算，如 5 MOD 2=1
\	取商运算，如 5\2=2	SQR（x）	求算术平方根，如 SQR（9）=3
ABS（x）	求绝对值，如 ABS（−4）=4	INT（x）	取整函数，如 INT（6.5）=6

4. 对输入语句的理解

输入语句的作用是实现算法的输入信息功能. 变量是指程序在运行时其值是可以变化的量，如 a，b，c 等都是变量，通俗把一个变量比喻成一个盒子（其实是内存单元），盒子内可以存放数据，可随时更新盒子内的数据.

5. 对输出语句的理解

输出语句的作用是实现算法的输出结果的功能，可以在计算机的屏幕上输出常量、变量的值等.

6. 对赋值语句的理解

赋值语句的作用是将表达式所代表的值赋给变量. 赋值语句中的“=”叫作赋值号，它和数学中的等号不完全一样；赋值号的左右两边不能对换，赋值语句是将赋值号右边的表达式的值赋给赋值号左边的变量，如 $a=b$ 表示用 b 的值代替变量 a 原先的值. 右边“表达式”可以是一个数据、常量和算式，如果“表达式”是一个算式时，赋值语句的作用是先计算出“=”右边表达式的值，然后将该值赋给“=”左边的变量，如若 $a=1$，$b=2$，$c=a+b$ 是指先计算 $a+b$ 的值 3 赋给 c，而不是将 $a+b$ 赋给 c.

输入语句和输出语句的区别在于：输入语句是外部直接给程序中变量赋值；输出语句是程序运行的结果输出到外部，先计算表达式，得到结果输出.

输入语句和赋值语句的区别在于：输入语句是外部直接给程序中变量赋值；赋值语句是程序内部运行时给变量赋值，先计算右边的表达式，得到的值赋给左边的变量.

教学建议

输入、输出语句与赋值语句都比较简单，容易掌握，可让学生多自主学习，上机操作，鼓励学生用这些语句来解决顺序结构的问题，特别提醒学生注意赋值语句的含义.

教学中，要注意将程序框图中的基本结构与伪代码相对应，在比较中理解，同时又关注伪代码的特殊性，必须遵循一些通用的格式规范. 要提醒学生注意阅读教材中的旁注信息，思考提出的每一个问题.

例 1 的教学重点在于熟悉三种语句的格式规范，特别要注意算式“$0.3x+y$”在伪代码中书写成“$0.3*x+y$”，是因为乘法运算符在伪代码中是用“$*$”表示的.

例 2 的教学重点在于帮助学生正确理解变量. 可以把变量看作一个存放数据的盒子，而且是最多只能存放一个数据的盒子. 当一个新数据放进去时，原来的数就被“挤”了出去. 代码第一行中 a 的值为 1，第二行“$a=a\wedge3+2$”表示将右边 a^3+2 的值 3 赋给 a，此时 a 的值变成 3，原来的值 1 就被冲掉.

例 3 中变量 a 的作用在于“过渡”，用于暂存数据. 第三行“$a=x$”将变量 x 的值赋给了 a，变量 x 空出；第四行“$x=y$”将变量 y 的值赋给了变量 x，变量 y 又空出；紧接着第五行“$y=a$”又将变量 a 的值赋给了变量 y，实现了两个变量 x，y 的值的交换.

例题分析

例. 阅读下面的伪代码，若输入 $x=2$，$y=4$，那么两输出的结果分别是什么？

```
INPUT  x，y
PRINT  x，y
x=y
y=x
PRINT  x，y
END
```

解 输入 $x=2$，$y=4$，第二行输出的结果仍是“2，4”，第三行“$x=y$”是把 y 的值 4 赋给了 x，也就是此时 x 的值变成了 4；第四行“$y=x$”是把 x 的值 4 赋给 y，所以最后输出的是“4，4”.

教学中可以让学生比较教材例 3 和上面的代码，体会要实现两个变量 x，y 的值的交换，中间变量是不可缺少的，进一步体会该算法的思想.

相关链接

程序设计方法简介

程序设计的基本目标是用算法对问题的原始数据进行处理，从而获得所期望的效果，但这仅仅是程序设计的基本要求. 要全面提高程序的质量，提高编程效率，使程序具有良好的可读性、可靠性、可维护性以及良好的结构，编制出好的程序来，应当是每位程序设计工作者追求的目标. 而要做到这一点，就必须掌握正确的程序设计方法和技术.

在程序设计初期，由于计算机硬件条件的限制，运算速度与存储空间都迫使程序员追求高效率，编写程序成为一种技巧与艺术，而程序的可理解性、可扩充性等因素被放到第二位.

随着计算机硬件与通信技术的发展，计算机应用领域越来越广泛，应用规模也越来越大，

程序设计不再是一两个程序员可以完成的任务．在这种情况下，编写程序不再片面追求高效率，而是综合考虑程序的可靠性、可扩充性、可重用性和可理解性等因素．正是这种需求刺激了程序设计方法与程序设计语言的发展．

1．早期程序设计

早期出现的高级程序设计语言有 FORTRAN、COBOL、ALGOL、BASIC 等语言．这一时期，由于追求程序的高效率，程序员过分依赖技巧与天分，不太注重所编写程序的结构，这时期可以说是无固定程序设计方法的时期．存在的一个典型问题是程序中的控制随意跳转，即不加限制地使用 goto 语句，这样的程序对别人来说是难以理解的，程序员自己也难以修改程序．

2．结构化程序设计

随着程序规模与复杂性的不断增长，人们也在不断探索新的程序设计方法，证明了只用三种基本的控制结构（顺序结构、条件结构、循环结构）即可实现任何单入口、单出口的程序．

Dijkstra 建议从一切高级语言中取消 goto 语句；Mills 提出程序应该只有一个入口和一个出口．这些工作导致了结构化程序设计方法的诞生．

而 Pascal 语言则是由 Niklaus Wirth 根据结构化程序设计方法开发出来的语言．其特点是提炼出程序设计共同的特征并能将这些特征编译成高效的代码，因而成为结构化程序设计的有力工具．C 语言也是一种广为流行的结构化程序设计语言，它具有灵活方便、目标代码效率高、可移植性好等优点．

3．面向对象程序设计

面向对象方法是一种把面向对象的思想运用于软件开发过程中，指导开发活动的系统方法，是建立在“对象”概念（对象、类和继承）基础上的方法学．对象是由数据和允许的操作组成的封装体，与客观实体有直接的对应关系．

面向对象的方法起源于面向对象的编程语言．20 世纪 60 年代后期就出现了类和对象的概念，类作为语言机制用来封装数据和相关操作．70 年代前期，Smalltalk 语言，奠定了面向对象程序设计的基础，1980 年 Smalltalk－80 标志着面向对象的程序设计已进入实用阶段．进入 80 年代相继出现了一系列面向对象的编程语言，如 C＋＋等．自 80 年代中期到 90 年代，面向对象的研究重点已经从语言转移到设计方法学方面，尽管还不成熟，但已陆续提出了一些面向对象的开发方法和设计技术．

11.3.2 条件语句

教材线索

本节从程序框图中的条件结构出发，引出条件语句的概念，指出条件语句即对应于程序框图中的判断框，并对照条件结构的一般形式给出条件语句的一般格式. 然后通过三个具体的实例逐步由简单到复杂，拓展了条件语句的范畴，引入了复合条件语句.

教学目标

1. 理解条件语句的功能和一般格式；能读懂包含条件的算法伪代码；
2. 能准确运用条件语句表达解决具体问题的算法；
3. 在具体问题的解决过程中，培养学生逻辑思维能力与表达能力，进一步体会算法思想.

教材分析

1. 重点

(1) 条件语句的格式和用法；

(2) 用自然语言、程序框图描述的包含条件结构的算法转化为伪代码的过程.

2. 难点

(1) 将具体问题的条件结构框图转化为条件语句的过程；

(2) 复合条件语句中的逻辑关系的理解.

3. 条件语句

条件语句有“IF—THEN—ELSE”和“IF—THEN”两种形式. “IF”表示条件语句的开始，“END IF”表示条件语句的结束.

第一种形式中，当计算机执行时，首先对 IF 后的条件进行判断，如果条件符合，就执行 THEN 后的语句，否则执行 ELSE 后的语句. 书写时一个条件语句中的 IF、ELSE 与 END IF 要对齐，内部语句体缩进且对齐，要掌握良好的书写风格.

第二种形式中，当计算机执行时，首先对 IF 后的条件进行判断，如果条件符合，就执行 THEN 后的语句，否则就直接结束该条件语句.

注意多个条件的语句表达方法：如 $a+b>c$ AND $b+c>a$ AND $a+c>b$. 在伪代码中，表达条件时往往会要用到一些关系运算符和逻辑运算符，常见的有：

运算符	功能说明
$<$	小于，与通常写法相同

$>$	大于，与通常写法相同
$<=$	小于或等于，相当于“$\leqslant$”，如 $x<=3$ 即 $x\leqslant3$
$>=$	大于或等于，相当于“$\geqslant$”，如 $x>=3$ 即 $x\geqslant3$
$<>$	不等于，相当于“$\neq$”，如 $x<>3$ 即 $x\neq3$
AND	且运算，表示 AND 前后两个条件必须同时成立才符合条件，相当于“且”
OR	或运算，表示 OR 前后两个条件只需要一个成立就符合条件，相当于“或”

条件语句的嵌套时，注意 END IF 是和最接近的匹配，要一层套一层，不能交叉.

教学建议

伪代码中的条件语句与程序框图中的条件结构存在一一对应关系，教学时应抓住条件结构在处理判定条件时，可能会出现多种情况，要尽量引导学生去思考这些情况，要求学生要考虑周全，不要遗漏.

安排例 1 的目的，不仅是为了应用条件语句，而且再次为学生提供了完整经历算法设计的全过程的机会. 代码并不复杂，不仅使学生进一步体会赋值语句、条件语句，而且还能锻炼学生阅读代码的能力.

例 2 是一道典型的可用条件结构的算法应用问题. 教学时，要求先用自然语言写出算法步骤，接着画出程序框图，最后把程序框图转化为伪代码. 由于这种应用问题实际上是分段函数模型，在前面已经接触过，需要用到条件结构的嵌套，弄清楚代码与程序框图之间的对应关系是正确书写伪代码的关键.

教材中提出的思考问题的伪代码如下：

```
INPUT a, b, c
d=b^2-4*a*c
IF d<0 THEN
    PRINT “方程没有实数根”
ELSE
    IF d=0 THEN
        PRINT “方程有两个相等实数根”
    ELSE
        PRINT “方程有两个不相等实数根”
    END IF
END IF
END
```

安排例 3 的目的一方面是进一步提高阅读能力，理解条件语句和赋值语句的完美结合，另一方面也是要与例 2 的条件语句相区别. 例 2 的两个条件语句形成了嵌套关系，例 3 中的三个条件语句是并列关系，没有形成嵌套.

例题分析

例. 编写一个伪代码，求一元二次方程 $ax^2+bx+c=0$（$a\neq0$）的根.

算法分析：计算 $\Delta=b^2-4ac$ 的值；

当 $\Delta>0$ 时，方程有两个不相等的实根 $x_1=\dfrac{-b+\sqrt{\Delta}}{2a}$，$x_2=\dfrac{-b-\sqrt{\Delta}}{2a}$；

当 $\Delta=0$ 时，方程有两个相等的实根 $x_1=x_2=\dfrac{-b}{2a}$；

当 $\Delta<0$ 时，方程没有实根.

程序框图如图 11－8 所示.

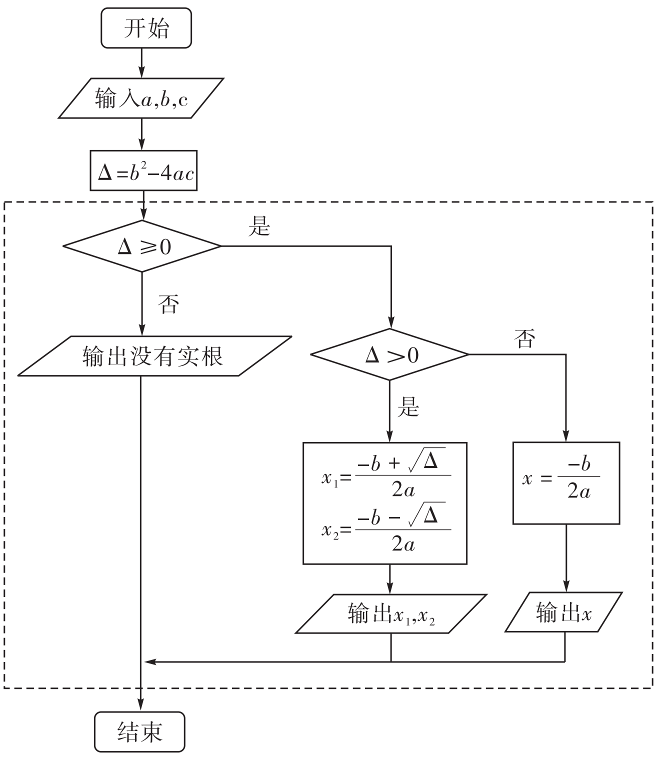


图 11－8

由程序框图写出相应的伪代码如下：

```
INPUT a,b,c
D=b^2-4*a*c
IF D>=0   THEN
      IF   D>0   THEN
            x1=(-b+SQR(D))/(2*a)
            x2=(-b-SQR(D))/(2*a)
            PRINT x1,x2
      ELSE
            x=-b/(2*a)
            PRINT x
      END IF
ELSE
     PRINT “没有实数根！”
END IF
END
```

相关链接

算法大师

· Rebort Sedgewick9b

算法的讲解者——RobertSedgewick 是普林斯顿大学的计算机科学教授. 他还是 Adobe Systems的一名主管，也曾作为访问学者在 XeroxPARC、IDA 和 INRIA 工作. 他在斯坦福大学获得博士学位 . 他的著作包括 Algorithm in C、Algorithm in C＋＋、Algorithm in Java 等系列书籍，这些都再版多次. “没有人能够将算法和数据结构解释得比 Robert Sedgewick 更清楚易懂了!”很多读过他著作的程序员这样说. 目前 Robert 正在研究算法设计、数据结构、算法分析等方面的基础理论. 他善于通过数学方法评估和预测算法性能，设法发现算法、数据结构的通用机制，例如使用逼近方法寻找更快速更高效的算法. 另外，他还将算法和图形学结合起来，例如使用可视化方法评估算法效率，算法的图形化模拟，用于出版物的高质量算法表现方法等等.

· Tony Hoare

计算机领域的爵士——TonyHoare，1934 年出生于英国，1959 年博士毕业于莫斯科国立大学，获得语言机器翻译专业学士学位. 1960 年发布了使他闻名于世的快速排序算法（Quick Sort），这个算法也是当前世界上使用最广泛的算法之一. Tony Hoare 在取得博士学位后，就职于 Elliott Brothers，领导了 Algo60 第一个商用编译器的设计与开发，由于其出色的成绩，最终成为该公司首席科学家. 从 1977 年开始，Tony Hoare 博士任职于牛津大学，投身于计算系统的精确性的研究、设计及开发. 因其对 Algoj 160 程序设计语言理论、互动式系统及 APL 的贡献，1980 年被美国计算机协会授予“图灵奖”. 1999 年在牛津大学退学后，Tony Hoare 博士被微软剑桥研究院聘请担任高级程序员，从事微软剑桥研究院研究生成果的工业化应用的工作，以及协助其它研究人员进行服务于软件产业及用户的长期基础研究项目. 2000 年因为其在计算机科学与教育上做出的贡献被封为爵士.

· Udi Manber

首席算法官——世界上还有如此奇怪的职位? 但是对于 Amazon 乃至 Google 来说，这一点也不奇怪. Udi Manber，这位前 Amazon 的“首席算法官”，现在是 Google 负责工程事务的副总裁. 他研究 WWW 的应用程序、搜索以及隐藏在这背后的算法设计. 在此期间，他与其他人共同开发了 Agrep、Glimpse 和 Harvest 等 Unix 上的搜索软件. 1998 年，Udi 成为了 Yahoo! 的首席科学家. 2002 年，Amazon 创造性地给了 Udi“首席算法官”的职位，和 Udi 为 Amazon 的“Search Inside the Book”搜索项目所做的工作相得益彰. Udi 还因为他所著的 Introduction to Algorithms——A Creative Approach 而被大家称道.

11.3.3 循环语句

教材线索

接上节，本节通过顺序结构用的输入、输出语句和赋值语句，条件结构的条件语句，从而提出循环结构的实现语句——循环语句. 并通过对应于“当型循环”和“直到型循环”提出了“WHILE 循环语句”和“UNTIL 循环语句”的一般形式，最后通过具体实例详细介绍循环语句的用法.

教学目标

1. 理解循环语句的两种形式，能读懂包含循环语句的算法代码；
2. 能正确运用循环语句表达解决具体问题的过程；
3. 通过具体实例算法设计，培养学生逻辑思维能力与表达能力，进一步体会算法思想.

教材分析

1. 重点

(1) 循环语句的表示方法、结构和用法；

(2) 由程序框图转化为程序语句时，条件结构和循环结构的区别.

2. 难点

(1) 将具体问题的程序框图转化为程序语句的过程；

(2) 当型循环和直到型循环在格式、逻辑方面的区别与联系.

3. 当型（WHILE 型）语句

当型循环语句由“WHILE—WEND”构成，当计算机遇到 WHILE 语句时，先判断条件的真假，如果条件符合，就执行 WHILE 与 WEND 之间的循环体；然后再检查上述条件，如果条件仍符合，再次执行循环体，这个过程反复进行，直到某一次条件不符合为止. 这时，计算机将不执行循环体，直接跳到 WEND 语句后，接着执行 WEND 之后的语句. 因此，当型循环有时也称为“前测试型”循环.

4. 直到型（UNTIL 型）语句

直到型循环语句由“DO—LOOP UNTIL”构成，当计算机遇到 DO 语句时，先执行 DO 和

UNTIL 之间的循环体，然后判断条件是否成立，如果不成立，再执行循环体. 这个过程反复执行，直到某一次符合条件为止，这时不再执行循环体，跳出循环执行 UNTIL 后面的语句. 因此，直到型循环有时也称为“后测试型”循环.

教学建议

循环语句是算法中的难点，也是重点，设计算法时一不小心容易造成死循环. 建议教师在充分理解循环语句的基础上，多通过例子分析，引导学生注意循环的入口数据和出口的条件，并能正确区分两种类型的循环.

教材中用不同类型的循环语句描述相同的算法，旨在说明两种循环语句是可以相互转化的. 循环语句中必定包含改变判断条件的语句，要特别注意循环控制条件在两种循环语句中的变化. 如用 WHILE 循环语句中的控制条件是“$k<=100$”，而改用 UNTIL 循环语句后的循环控制条件就变成了“$k>100$”.

例 1 用到的是 UNTIL 循环语句，如果改用 WHILE 循环语句，代码如下：

```
k=1
WHILE k≤10
    INPUT x
    y=x^3-25*x+7
    PRINT y
    k=k+1
WEND
END
```

例 2 的难点在于循环体中变量的控制技巧，教学时，除要求学生懂得循环语句的作用外，还可以向学生介绍这种节约变量的意义.

例 3 的算法综合了三种基本逻辑结构，特别是条件结构与循环结构的配合使用，教学时，应该让学生经历算法思路分析、用自然语言描述算法步骤、绘制程序框图、编写伪代码的全部过程，

例题分析

例猴子第一天摘下若干个桃子，当即吃了一半，还不过瘾又多吃了一个；第二天早上又将剩下的桃子吃了一半又多吃了一个；以后每天早上都吃了前一天的一半又多一个，到第 10 天早上想再吃时，见只剩下一个桃子，试编写一个程序，求第一天共摘了多少桃子？

算法分析：设第一天、第二天、第三天、……、第十天的桃子数分别为 s_1，s_2，s_3，…，s_{10}，则 $s_{10}=1$，$s_9=2\times(s_{10}+1)$，$s_8=2(s_9+1)$，…，$s_2=2\times(s_3+1)$，$s_1=2\times(s_2+1)$.

直接写出伪代码如下：

```
s=1
i=1
WHILE i<10
    s=2*(s+1)
    i=i+1
WEND
PRINT s
END
```

相关链接

计算机科学和数学的关系

计算机科学和数学的关系有点奇怪．二三十年以前，计算机科学基本上还是数学的一个分支，而现在，计算机科学拥有广泛的研究领域和众多的研究人员，在很多方面反过来推动数学发展，从某种意义上可以说是孩子长得比妈妈还高了．但不管怎么样，这个孩子身上始终流着母亲的血液．这血液是 the mathematical underpinning of computer science（计算机科学的数学基础）——也就是理论计算机科学．

现代计算机科学和数学的另一个交叉是计算数学/数值分析/科学计算，传统上不包含在理论计算机科学以内．所以本文对计算数学全部予以忽略．最常和理论计算机科学放在一起的一个词是什么？离散数学！这两者的关系是如此密切，以至于它们在不少场合下成为同义词．

传统上，数学是以分析为中心的．数学系的同学要学习三四个学期的数学分析，然后是复变函数，实变函数，泛函分析等等．实变函数和泛函分析被很多人认为是现代数学的入门．在物理，化学，工程上应用的，也以分析为主．

随着计算机科学的出现，一些以前不太受到重视的数学分支突然重要起来．人们发现，这些分支处理的数学对象与传统的分析有明显的区别：传统的分析研究对象是连续的，因而微分，积分成为基本的运算；而这些分支研究的对象是离散的，因而很少有机会进行此类的计算．人们从而称这些分支为“离散数学”．“离散数学”的名字越来越响亮，最后导致以分析为中心的传统数学分支被相对称为“连续数学”．

11.4　算法案例

教材线索

本节通过大量的算法案例：辗转相除法、韩信点兵问题（即中国剩余定理）、二分法求根、秦九韶算法等重要的算法例子，引导学生学习其中的数学算法思想，全面提高学生的数学素养和算法设计能力.

教学目标

1. 理解辗转相除法、中国剩余定理、二分法、秦九韶算法；

2. 通过阅读中国古代中的算法案例，体会中国古代数学对世界数学发展的贡献；

3. 通过经典算法案例的学习，培养学生抽象概括能力、逻辑思维能力与语言表达能力，进一步体会算法思想.

教材分析

1. 重点

（1）以典型算法案例为载体，使学生通过模仿、操作、探索经历算法设计的全过程；

（2）感受算法在实际问题中的重要作用，进一步体会算法的基本思想.

2. 难点

（1）对几个典型算法案例的理解；

（2）提炼实际问题的算法中的循环结构，并用程序框图和伪代码表示出来.

3. 辗转相除法

辗转相除法，又称欧几里得算法，是求两个正整数最大公约数的算法. 辗转相除法首次出现于欧几里得的《几何原本》（第Ⅶ卷，命题ⅰ和ⅱ）中，“设有不相等的二数，若依次从大数中减去小数，若余数总是量不尽它前面一个数，直到最后的余数为一个单位，则该二数互素.”而在中国则可以追溯至东汉出现的《九章算术》.

两个整数的最大公约数是能够同时整除它们的最大的正整数. 辗转相除法基于如下原理：两个整数的最大公约数等于其中较小的数和两数的差的最大公约数. 例如，252 和 105 的最大公约数是 21（252＝21×12；105＝21×5）；因为 252－105＝147，所以 147 和 105 的最大公约数也是 21. 在这个过程中，较大的数缩小了，所以继续进行同样的计算可以不断缩小这两个数直至其中一个变成零. 这时，所剩下的还没有变成零的数就是两数的最大公约数.

4. 中国剩余定理

中国剩余定理是中国古代求解一次同余式组的方法，是数论中一个重要定理，又称孙子定理. 孙子定理对近代数学如环论，赋值论都有重要影响.

公元前后的《孙子算经》中有“物不知数”问题：“今有物不知其数，三三数之余二，五五数之余三，七七数之余二，问物几何?”答为“23”. 也就是求同余式组 $x\equiv2$（mod 3），$x\equiv3$（mod 5），$x\equiv2$（mod 7）（式中 $a\equiv b$（mod m）表示 m 整除 $a-b$）的正整数解. 明朝程大位用歌谣给出了该题的解法：“三人同行七十稀，五树梅花廿一枝，七子团圆月正半，除百零五便得知.”即解为 $x\equiv2\times70+3\times21+2\times15\equiv233\equiv23$（mod 105）.

5. 二分法

通过每次把 f（x）的零点所在小区间收缩一半的方法，使区间的两个端点逐步迫近函数的零点，以求得零点的近似值，这种方法叫做二分法. 在前面的学习中已经接触过，由于计算过程的具体运算复杂，但每一步的方式相同，所以可通过循环结构来实现算法.

6. 秦九韶算法

秦九韶算法是中国南宋时期的数学家秦九韶提出的一种多项式简化运算的算法. 在西方被称作霍纳算法. 该算法看似简单，其最大的意义在于将求 n 次多项式的值转化为求 n 个一次多项式的值. 在人工计算时，利用秦九韶算法和其中的系数表可以大幅简化运算；对于计算机程序算法而言，加法比乘法的计算效率要高很多，因此该算法仍有极大的意义. 对于计算机来说，做一次乘法运算所用的时间比做一次加法运算要长得多，所以此算法极大地缩短了 CPU 运算时间.

教学建议

本节的例题较多，同时也比较难，对于每个例子都要详细分析其思想，使学生理解每个算法的思考过程，全面提高学生的逻辑思维能力.

关于案例 1 的教学，教师可以与之前介绍的“更相减损术”比较，通过具体的两个数理解算法思想. 教材是用直到型循环结构给出辗转相除法的算法步骤、程序框图和伪代码，随后要求学生改用当型循环结构来描述算法、程序框图和伪代码. 目的在于除让学生进一步通过比较加深对循环结构的认识外，更重要的是为学生设计算法、体会算法思想提供机会. 教学时应留时间让学生实践，引导学生得出自己设计的算法步骤、程序框图和伪代码.

当型循环结构的程序框图如图 11－9.

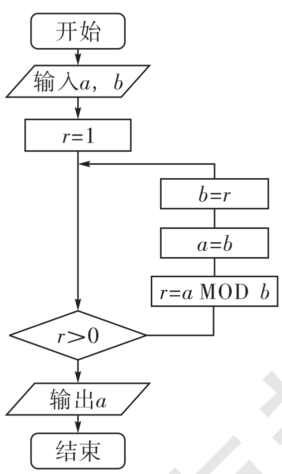


图 11－9

伪代码如下：

```
INPUT a, b
r=1
WHILE r>0
    r=a MOD b
    a=b
    b=r
WEND
PRINT a
END
```

关于案例 2 的教学，重点应放在问题的通用解法上，即构造不定方程组的算法分析. 该问题也可以用“韩信点兵”问题作为背景引入. 韩信带 1 500 名兵士打仗，战死四五百人，站 3 人一排，多出 2 人；站 5 人一排，多出 4 人；站 7 人一排，多出 6 人. 韩信马上说出人数：1 049.

案例 3 的理解相对更难一些，教学中应注意，当有解区间 $[a, b]$ 已经满足精度要求，即 $|b-a|<c$ 时，教材中是再次取区间 $[a, b]$ 的中点 x_0 作为近似解. 实际上，此时区间 $[a, b]$ 中的实数都是满足精度要求的近似解. 如果直接取区间端点作为近似解，程序框图可以简化为图 11－10 的形式.

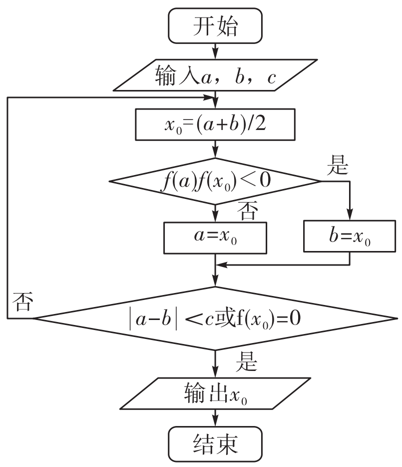


图 11－10

案例 4 的教学中，教材先安排求简单的一元多项式的值，并对两种算法的效率进行比较分析，目的在于使学生了解解决同一问题的算法可能有很多种. 但算法有“优”“劣”之分，判断标准之一就是运算的效率. 这里通过统计乘法和加法的次数也为后面讲秦九韶算法中多项式的变形作了铺垫.

关于运算效率问题不是我们要分析的重点. 教材也是通过具体的例子，阐述多项式的分解、转化求值来讲秦九韶算法，然后再归纳出一般性的算法. 找到递推关系“$v_k=v_{k-1}x+a_{n-k}$（$k=1, 2, \cdots, n$）”是实现循环结构的关键所在.

本节中的中国剩余定理、秦九韶算法都是我国古代数学中的精华，教学中要广泛收集素材，向学生介绍中国古代数学的特点、成就和对世界数学发展的贡献.

例题分析

为便于理解教材上的案例 3，可以补充下面的问题.

例. 在一个风雨交加的夜里，从某水库闸门到防洪指挥部的电话线路发生了故障，这是一条 10 km 长的电话线路，每隔 50 m 有一根电线杆，维修工人需爬上电线杆测试，如何迅速查出故障所在？如果沿着线路一小段一小段地查找，困难很多，每查个电线杆要爬一次电线杆呢.

想一想，你能帮他找到一个简单易行的方法吗？

思路引导：如图 11－11 所示，他首先从中点 C 开始查，用随身带的话机向两端测试时，发现 AC 段正常，断定故障在 BC 段，再到 BC 中点 D，这次发现 BD 段正常，可见故障在 CD 段，再到 CD 段中点 E 来查.

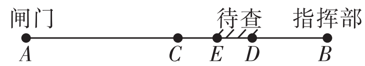


图 11－11

每查一次，可以把待查的线路长度缩减一半，算一算，要把故障可能发生的范围缩小到 50 m至 100 m 左右，总共需要几次？

实际上只需经过七次即可.（第一次：5 km，第二次：2.5 km，第三次：1.25 km，第四次：625 m，第五次：312 m，第六次：156 m，第七次：78 m）

这是采取逐步缩小范围的办法找故障，当范围越小越容易查找发生故障的电话线路. 在数学中我们也可以用这样的方法来解方程.

相关链接

古中国人在筹算、观天和算法方面的成就

我国是世界上最早的文明国家之一. 很早以前，我们的祖先在渔猎农事活动中就接触到了计算和测量，并在这方面积累了大量的知识.

万里长城和大运河是我国古代文明的伟大成就. 战国时期战争连绵，燕、赵、秦三国为了抵御来自北方的侵扰，建筑了长城；秦始皇统一全国，把它们连接起来. 后来，汉朝和明朝都大规模修筑过长城. 长城由西至东，在险峻起伏的山岭上绵延数千公里，是世界上仅有的巨大土石建筑. 沟通南北的大运河，长达一千七百多公里，朴实壮观，是非常杰出的水利工程. 我国人民在长城和运河的建造过程中积累了大量的几何测量、数字计算和土木工程方面的知识.

我国古代的计算不是用记数文字直接进行，而是用算筹，很有特色. 在开始的时候，人们是用一些小树枝来计数，一根小树枝代表一头牲畜、一堆谷物或者一件农具. 后来，逐渐形成了一套计算方法，小树枝也慢慢变成了竹制、铁制、牙制的小棍，外形规格齐整，这就是算筹.

筹算可以进行整数和分数的加、减、乘、除、开方等各种运算. 直到元、明以前，筹算一

直是我国的主要计算方法.

筹算的记数法既是十进，又按位值分别表示不同单位，和现代记数法相似. 著名的数学著作《九章算术》，大约编于公元四、五十年间的东汉初期. 这部书是采用问题集的形式编的，共有二百四十六个问题，分成方田、粟米、衰分、少广、商功、均输、盈不足、方程和勾股九章.

方田章讲的是各种分数计算和方田、梯形田、斜方形田、圆田、半圆形田、弧田、环形田等的面积计算；粟米章讲的是粮食交易的简单比例计算；衰分章讲的是一些按比例分配的问题；少广章讲的是由已知面积和体积，反求边的长短和面的宽广的问题，其中总结出了开平方和开立方的方法；商功章讲的是计算各种体积的方法，主要解决筑城、建堤、挖沟、修渠等实际工程问题；均输章讲的是粮食运输均匀负担的计算方法；盈不足章讲的是盈亏计算法和它的应用；方程章讲的是正负数算法，还有各种三元一次和四元一次联立方程的解法. 勾股章叙述了勾方、股方的和等于弦方的勾股定理，以及相似直角三角形解法的问题.

《九章算术》的内容丰富多彩，包括了许多算术、几何、代数和三角的知识，是一部非常杰出的数学专著，它对我国数学的发展影响深远.

《九章算术》不只在中国数学史上占有十分重要的地位，而且影响远及国外. 朝鲜和日本都曾经用它作为教科书. 欧洲在中世纪的一些算法，比如分数和比例就很可能是从中国传入印度、再经阿拉伯传入欧洲的. 在阿拉伯和欧洲的早期数学著作中，把“盈不足”称为“中国算法”就是一个证明. 现在,《九章算术》已作为世界科学名著，被译成许多种文字出版.

《周髀算经》是我国另一部有名的天文学、数学著作，大约时在公元前一百年前后的西汉年间成书. 书里明确给出了勾股定理的一般形式，即勾^2＋股^2＝弦^2.

书中介绍了在两地利用标杆测出日影、再进一步利用勾股定理，算出太阳高度的方法，即书中还谈到了用一根直径一寸、长八只的中空竹管观测太阳，太阳的圆影正好与竹管的视线吻合，再进一步利用勾股定理推算出太阳的直径来. 这说明我们的祖先至少在西汉年间，就能正确地应用直角三角形的勾股定理了.

等到三国时代，吴国人赵爽用几何方法对勾股定理进行了相当严格的论证. 公元前五百年，春秋战国时代的学者已经有了相当丰富的数学知识. 庄子《天下篇》中有“一尺之捶，日取其半，万世不竭”的记载. 意思是一根一尺长的木棍，每天截掉一半，千年万载也截不完. 直到今天，人们还常把“日取其半”作为了解“极限”思想的典型例子.

大约在四千五百到三千五百年前的这段时期里，我国发明了第一辆车子. 另外，从我国出土的许多殷代以前的陶器上也能看到不少圆形图案. 这说明很早以前，我们的祖先就认识圆了.

在《周辨算经》周公和商高的对话中，谈到“周三径一”，这是我国最初的圆周率，被称为古率. 后来，圆周率数值的精确性不断得到提高.

我国最早用严密的数学方法来求算圆周率数值的是刘徽. 他认为古率为 3，是圆内接正六边形的周长对直径的比值，这比圆周长对直径的比值要小得多.

刘徽把圆内接正六边形各边所对的弧平分，做出圆内接正十二边形，利用勾股定理求出它的边长. 同理，可以求出圆内接正二十四、四十八、九十六边形的边长. 内接正多边形的边数

越多，求出的圆周率数值也就越准确．这就是刘徽的“割圆术”．

“割圆术”用折线逐步逼近曲线，用圆内接正多边形的面积逐步逼近圆面积，这种用有限来逼近无限的方法，不仅提供了比较精确的圆周率的数值，而且为后来计算圆周率的人们奠定了坚实可靠的理论基础．

——选自王利公《数学的童年》

习题参考答案

习题 1

1. 算法步骤如下：

S1：用 2 除 7，得到余数 1，所以 2 不能整除 7；

S2：用 3 除 7，得到余数 1，所以 3 不能整除 7；

S3：用 4 除 7，得到余数 3，所以 4 不能整除 7；

S4：用 5 除 7，得到余数 2，所以 5 不能整除 7；

S5：用 6 除 7，得到余数 1，所以 6 不能整除 7. 因此 7 是质数.

2. 操作步骤如下：

步骤	a	b
1	147	273
2	147	126（=273−147）
3	21（=147−126）	126
4	21	105（=126−21）
5	21	84（=105−21）
6	21	63（=84−21）
7	21	42（=63−21）
8	21	21（=42−21）

所以，147 和 273 的最大公约数是 21.

3. 该算法的作用是求函数 $f(x)=\begin{cases}2x,\ (x\geqslant 1),\\ x^2+1,\ (0\leqslant x<1)\end{cases}$ 的函数值，所以当输入 $x=0$ 时，输出的函数值最小为 1.

11.2.1　练习

1. 当 $x=3$ 时，$y=7$，$b=7^2+1=50$，即输出结果是 50.

2. (1) 当 $a=2$ 时，$x=2^2=4$，$y=4^2+4+1=21$，即输出值是 21；

(2) 当输出值 $y=3$ 时，即 $x^2+x+1=3$，$x=1$ 或 $x=-2$，所以只能是 $a^2=1$，从而输入的值 $a=1$ 或 $a=-1$.

3. 算法步骤如下：

S1：①+②×2 得 $5x=3$；

S2：解 $5x=3$ 得 $x=\frac{3}{5}$；

S3：②－①×2 得 $5y=-1$；

S4：解 $5y=-1$ 得 $y=-\frac{1}{5}$.

程序框图如图 11－12 所示.

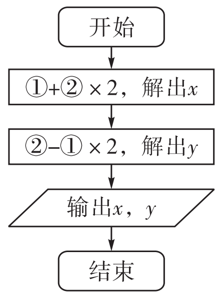


图 11－12

4. 算法步骤如下：

S1：输入三角形三边长 a，b，c；

S2：计算 $p=\frac{a+b+c}{2}$；

S3：计算 $S=\sqrt{p(p-a)(p-b)(p-c)}$；

S4：输出 S.

程序框图如图 11－13 所示.

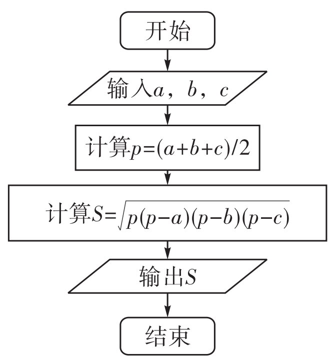


图 11－13

11.2.2　练习

1. 若 $x=5$，则输出 $y=3\times5^2+1=76$；若 $x=2$，则输出 $y=2^2-1=3$.

2. 算法步骤如下：

S1：将九个玉镯分成三组，每组三个，任取两组置于天平左右两边称量；

S2：若天平平衡，则假玉镯在另一组中；若天平不平衡，则假玉镯在较重的一组中；

S3：在包含假玉镯的一组中任取两个称量；

S4：若天平平衡，则假玉镯是另一个；若天平不平衡，则假玉镯是较重的那个.

3. 算法步骤如下：

S1：输入自变量 x；

S2：判断若 $x\geqslant 0$，则输出 $f(x)=1$；否则，输出 $f(x)=-1$.

程序框图如图 11－14 所示.

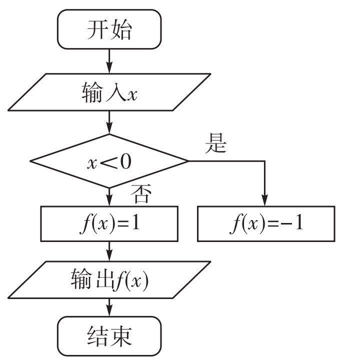


图 11－14

4. 算法步骤如下：

S1：输入原始分数 x；

S2：判断若 $x\geqslant 60$，则等级 $D=$“合格”；否则，等级 $D=$“不合格”；

S3：输出等级 D.

程序框图如图 11－15 所示.

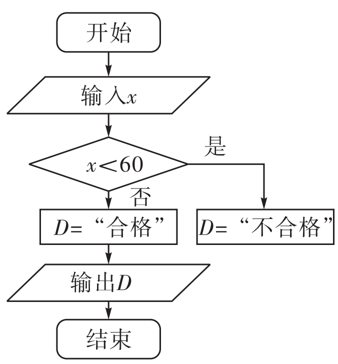


图 11－15

11.2.3 练习

1. 算法步骤如下：

S1：赋初始值 $k=1$，$T=1$；

S2：$T=Tk$，$k=k+2$；

S3：判断，若 $k>99$，则执行 S4，否则返回 S2；

S4：输出 T.

程序框图如图 11－16 所示.

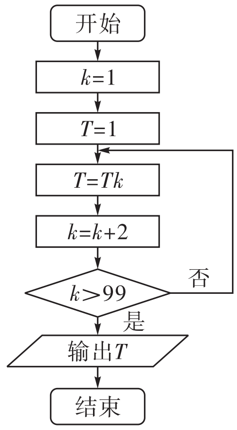


图 11－16

2. 算法步骤如下：

S1：赋初始值 $k=1$，$S=0$；

S2：$S=S+k^2$，$k=k+1$；

S3：判断，若 $k>100$，则执行 S4，否则返回 S2；

S4：输出 S.

程序框图如图 11－17 所示.

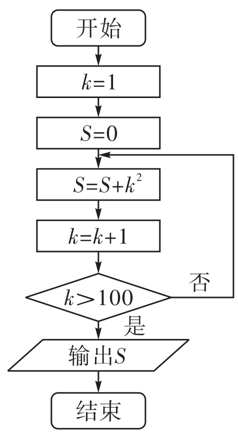


图 11－17

3. 由 $2+4+6+\cdots+2k=2\ 070$ 知，$k=45$，即 $2+4+6+\cdots+90=2\ 070$，所以“?”处填 91 或 92 均可以满足要求.

4. 依次输出 3×1，3×2，3×3，$\cdots$，3×33，即输出 100 以内能被 3 整除的正整数.

习题 2

1. C. 算法通常是指由有限多个步骤组成的求解某一类问题的通用的方法，对于该类问题中

的每个给定的具体问题，机械地执行这些步骤就可以得到问题的解答．算法有三个特征：确定性、有效性和有限性．选项 C 不是算法．故选 C.

2. A. 根据算法的特点：如果在执行过程中，不需要分类讨论，则不需要有条件结构；如果不需要重复执行某些操作，则不需要循环结构；但任何一个算法都必须有顺序结构．故选 A.

3. 算法步骤如下：

S1：初始变量 $k=1$；

S2：输出 $a=15k$，$k=k+1$；

S3：判断，若 $k\leqslant 66$，则返回 S2；否则结束算法.

程序框图如图 11－18 所示.

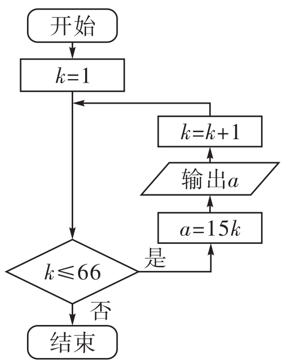


图 11－18

4. 算法步骤如下：

S1：输入自变量 x；

S2：若 $x\leqslant 1$，则 $y=x$；若 $1<x<10$，则 $y=2x-1$；若 $x\geqslant 10$，则 $y=3x-11$；

S3：输出函数值 y.

程序框图如图 11－19 所示.

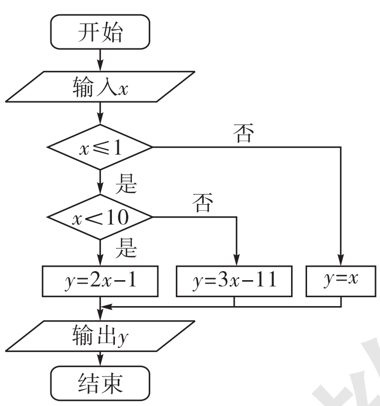


图 11－19

5. 算法步骤如下：

S1：输入楼层 x；

S2：判断，若 $x\leqslant 3$，则 $y=10$；否则，$y=10+2(x-3)$；

S3：输出电梯运行费 y.

程序框图如图 11－20 所示.

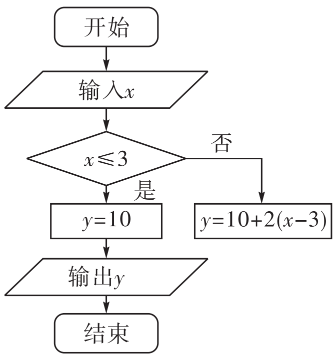


图 11－20

6. 判断框内的条件是"$k\geqslant 5$".

7. 算法功能是计算 $100\times 99\times 98\times\cdots\times 2\times 1$ 的结果.

11.3.1 练习

1. 算法步骤如下：

S1：输入圆的半径 r；

S2：计算 $C=2\times 3.14r$，$S=3.14r^2$；

S3：输出 C，S.

程序框图如图 11－21 所示.

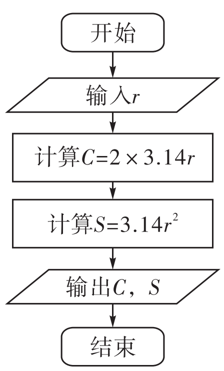


图 11－21

伪代码（如图 11－22）：

```
INPUT “r=”;  r
C=2*3.14*r
S=3.14*r^2
PRINT P, S
END
```

图 11－22

2. 算法步骤如下：

S1：输入成绩 a，b，c；

S2：计算 $w=\dfrac{a+b+c}{3}$；

S3：输出 w.

程序框图如图 11－23 所示.

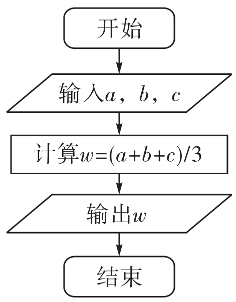


图 11－23

伪代码（如图 11－24）：

```
INPUT “a, b, c=” ; a, b, c
w=(a+b+c) / 3
PRINT w
END
```

图 11－24

3. $a=7$，$b=15$

4. 算法步骤如下：

S1：输入自变量 x；

S2：计算 $y=x^2+2x-3$；

S3：输出 y.

程序框图如图 11－25 所示.

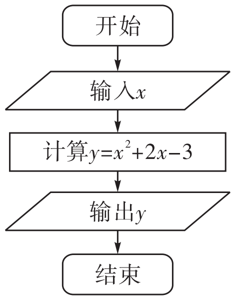


图 11－25

伪代码：

```
INPUT "x="; x
y=x^2+2*x-3
PRINT y
END
```

11.3.2 练习

1. 算法步骤如下：

S1：输入 x；

S2：若 $x\geqslant 0$，则 $|x|=x$；若 $x<0$，则 $|x|=-x$；

S3：输出 $|x|$.

程序框图如图 11－26 所示.

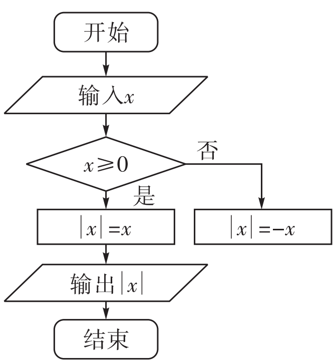


图 11－26

伪代码：

```
INPUT “x=”; x
IF x>=0 THEN
   y=x
ELSE
   y=-x
END IF
PRINT “|x|=”;  y
END
```

2. 算法步骤如下：

S1：输入自变量 x；

S2：若 $x<1$，则 $y=x+1$；若 $1\leqslant x<10$，则 $y=2x$；若 $x\geqslant 10$，则 $y=3x-10$；

S3：输出 y.

程序框图如图 11－27 所示.

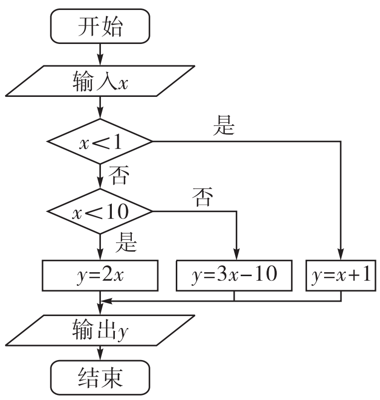


图 11－27

伪代码：

```
INPUT “x=” ; x
IF  x<1  THEN
   y=x+1
ELSE
   IF x<10 THEN
      y=2*x
   ELSE
      y=3*x-10
   END IF
END IF
PRINT y
END
```

3. 伪代码：

```
INPUT “汇款金额”； x
IF  x<=100  THEN
    y=1
ELSE
    IF  x≤5 000  THEN
        y=0.01*x
    ELSE
        y=50
    END  IF
END IF
PRINT  “手续费”；  y
END
```

4. 输出结果是−13，该算法的功能是找出输入的三个数中的最小的数并输出.

11.3.3　练习

1. 分别用 UNTIL 语句和 WHILE 语句设计的算法伪代码：

```
k=100
T=1
DO
    T=T*k
    k=k-1
LOOP  UNTIL  k<1
PRINT  T
END
```

```
k=100
T=1
WHILE  k>0
    T=T*k
    k=k-1
WEND
PRINT  T
END
```

2. 分别计算 $2+3+\cdots+21$ 和 $1\times2\times3\times\cdots\times21$ 的结果并输出.

3. 算法步骤如下：

S1：初始化变量 $k=1$，$S=0$；

S2：计算 $S=S+\dfrac{1}{k(k+1)}$，$k=k+1$；

S3：若 $k>99$，则执行 S4；否则返回 S2；

S4：输出 S.

程序框图如图 11－28 所示.

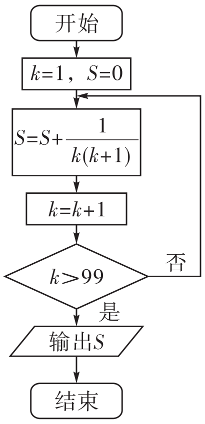


图 11－28

伪代码：

```
k=1
S=0
DO
    S=S+1 / (k*(k+1))
    k=k+1
LOOP  UNTIL  k>99
PRINT S
END
```

4．算法步骤如下：

S1：初始变量 $a=0$，$b=0$，$c=0$；

S2：输入分数 x；

S3：若 $x\geqslant 80$，则 $a=a+1$；若 $60\leqslant x<80$，则 $b=b+1$；若 $x<60$，则 $c=c+1$；

S4：判断输入的分数是否达到 40 个，若是，则执行 $S5$；若没有，则返回 $S2$；

S5：输出 a，b，c.

程序框图如图 11－29 所示.

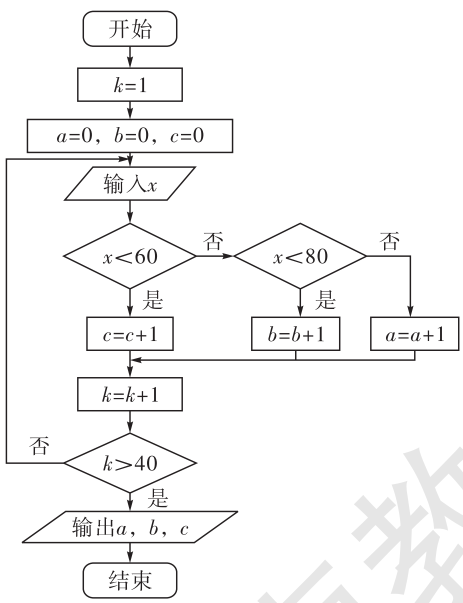


图 11－29

伪代码：

```
k=1
a=0
b=0
c=0
WHILE  k<=40
    INPUT  x
    IF x<60 THEN
       c=c+1
    ELSE
       IF  x<80  THEN
          b=b+1
       ELSE
          a=a+1
       END IF
    END IF
    k= +1
WEND
PRINT a, b, c
END
```

习题 3

1. $y=\begin{cases}0 & (x<0)\\ 1 & (0\leqslant x<1)\\ x & (x\geqslant 1)\end{cases}$

2. 伪代码：

```
INPUT  "F=";  x
y=(x-32)*5/9
PRINT  "C=";  y
END
```

3. 伪代码：

```
INPUT  "x=";  x
a=x MOD 2
IF  a=0  THEN
   PRINT  "x是偶数"
ELSE
   PRINT  "x是奇数"
END IF
END
```

4. 伪代码：

```
INPUT  m
k=2
DO
    a=m MOD k
    IF  a=0  THEN
        PRINT  "x不是质数"
    ELSE
        k=k+1
    END IF
LOOP UNTIL  a=0  OR  k>m-1
IF  k=m  THEN
    PRINT  "x是质数"
END IF
END
```

5. 伪代码：

```
INPUT  "t=";  t
IF  t<=1  THEN
    y=5
ELSE
    IF  t<6  THEN
        y=5+t-1
    ELSE
        y=10
    END IF
END IF
PRINT  y
END
```

6. 伪代码：

```
INPUT  "Year=";  y
a=y MOD 4
b=y MOD 100
c=y MOD 400
IF  a=0  AND  b<>0  THEN
    PRINT  "闰年"
ELSE
    IF c=0 THEN
        PRINT  "闰年"
    ELSE
        PRINT  "不是闰年"
    END IF
END IF
END
```

7. 算法步骤如下：

S1：初始化变量 $k=1$，$S=0$；

S2：计算 $S=S+(-1)^{k+1}\frac{1}{k}$，$k=k+1$；

S3：若 $k>100$，则执行 S4；否则返回 S2；

S4：输出 S.

程序框图如图 11－30 所示.

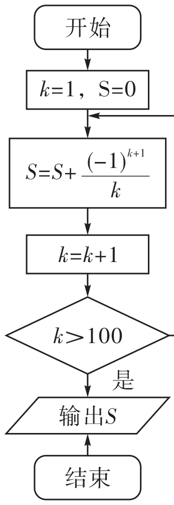


图 11－30

伪代码：

```
k=1
S=0
DO
      S=S+((−1)^(k+1))/k
      k=k+1
LOOP  UNTIL  k>100
PRINT  S
END
```

习题 4

1. ∵$204=85\times2+34$，

$85=34\times2+17$，

$34=17\times2$，

∴204 和 85 的最大公约数是 17.

2. 53

3. 约为 1.32 495

4. -1

5. 55

6. 389

复习题十一

1. 30

2. 22，−22

3. 2 881. 75

4. 程序框图如图 11－31 所示.

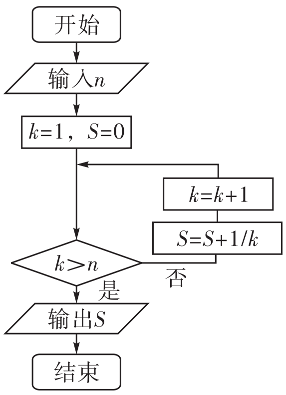


图 11－31

伪代码：

```
INPUT  n
k=1
S=0
DO
      S=S+1 / k
      k=k+1
LOOP  UNTIL  k>n
PRINT  S
END
```

5. 程序框图如图 11－32 所示.

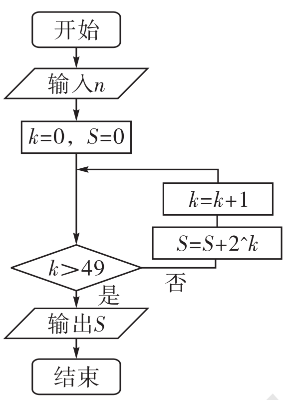


图 11－32

6. 7 只小兔，10 只鸡.

7. 算法伪代码如下：

```
n=2 009
a=200
WHILE a<=300
    a=a+0.05*a
    n=n+1
WEND
PRINT n
END
```

8. 程序框图如图 11－33 所示.

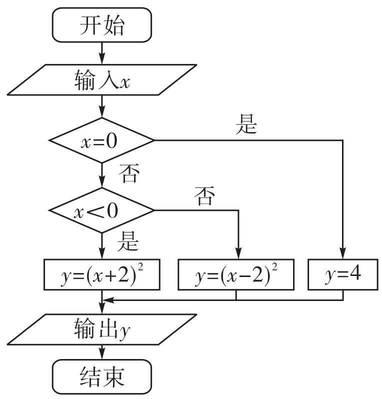


图 11－33

9. 57

10. 1. 416 016

11. 算法伪代码如下：

```
INPUT t
IF t<=3 THEN
  y=0.22
ELSE
  y=0.22+0.1*(t-3)
END IF
PRINT y
END
```

12. 提示：可以利用关键词搜索网络资源.

第 12 章　统计学初步

一、教学目标

1. 理解总体，个体和总体均值的含义，能在实例中加以分析.

2. 理解样本及样本均值的含义，理解它们的随机性.

3. 理解方差和标准差的含义，能区分总体方差与样本方差，了解样本方差的随机性.

4. 了解抽样调查的必要性和可能性，了解有放回抽样与无放回抽样的差异，理解简单随机抽样的含义，能在实际中合理地设计抽样方案.

5. 了解调查问卷的设计要求.

6. 理解分层抽样方法，了解系统抽样方法.

7. 理解频率的定义，能分析和绘制频率分布表，能根据经验公式和具体实际合理地选择分组.

8. 了解频率直方图，能分析和绘制频率分布直方图.

9. 了解频率折线图，能读懂频率折线图.

10. 了解数据的茎叶图与双茎叶图，能绘制简单的茎叶图.

11. 理解数据相关性的含义，能通过数据散点图初步判断数据间的相关性.

12. 理解回归直线的含义，能通过公式求出回归直线. 了解参数，最小二乘估计的含义. 理解参数未知性的特点. 了解如何对未知参数进行统计推断为统计学的根本任务之一.

二、教材说明

统计学是研究如何从数据中提取有用信息的科学，内容包括如何收集和分析数据.

统计学的思想远古即存，但作为一门学科的历史却不长. 直到二十世纪初，英国著名统计学家卡尔·皮尔逊. R. A. 费歇尔，美国著名统计学家奈曼等人才建立起现代统计学的基本框架，此后，统计学才得以蓬勃的发展. 在我国，1992 年 11 月国家技术监督局颁布的 GB/T14745—9《学科分类与代码》中，才将统计学与数学、经济学并列上升为一级学科.

教材通过大量贴近生活的实例引入统计学的一些基本概念，如总体，个体，样本，均值，方差等等，体现了统计学是一门应用性很强的学科. 教师讲授时既要注意统计思想的灌输，又要注意把统计学和纯数学区别开来，不要把统计学讲成一门数学，同时还要注意不能完全把统

计学和数学割裂开来，这是教学的一个难点问题.

本章主要内容有：总体、个体、样本、样本量的概念以及与之相联系的均值、方差、标准差的概念. 抽样调查的方法，抽样调查的必要性，抽样方案的设计，调查问卷的设计，分层抽样与系统抽样. 用样本分布估计总体分布. 频率分布表，频率分布直方图，数据的茎叶图. 最后介绍数据的相关性概念与简单的一元回归分析. 内容安排既贴近生活，体现了统计学应用性强的特点，又循序渐进，符合学生的认知特点.

教材十分注意统计思想的灌输. 由样本推断总体，窥一斑而知全豹. 统计学是一门方法论的学科，怎样正确运用统计方法解决实际问题尤为重要. 教材通过阅读材料“《文学摘要》的破产”来揭示了正确运用统计方法的重要性.

统计是与数据相联系的，而现实生活中数据常常是海量的，此时手工计算已经不可能. 教材最后介绍了如何使用计算机与统计软件来进行简单的计算. 在教学中，要鼓励学生使用现代信息技术手段来处理繁杂的计算，解决实际问题.

三、课时安排建议

本章教学约需 16 课时，具体分配如下（仅供参考）：

12.1　总体和个体

12.1.1	总体、个体和总体均值	1 课时
12.1.2	样本与样本均值	1 课时
12.1.3	方差与标准差	2 课时

12.2　抽样调查方法

12.2.1	随机抽样	1 课时
12.2.2	调查问卷的设计	1 课时
12.2.3	分层抽样和系统抽样	1 课时

12.3　用样本分布估计总体分布

12.3.1	频率分布表	1 课时
12.3.2	频率分布直方图	1 课时
12.3.3	频率折线图	1 课时
12.3.4	数据茎叶图	1 课时

12.4　数据的相关性

12.4.1	相关性	1 课时
12.4.2	回归直线	2 课时
	小节与复习	2 课时

四、教学建议

1. 本章内容是在初中“统计初步”的基础上学习的. 统计学是收集、整理和分析数据的学

科，应用面极广，工业、农林、经济、金融、医药、生物、管理、社会、国防等各个领域都经常使用各种统计方法．因此在本章的教学中，一定要强调与实际结合．

2．统计学是一门方法论的学科，是一门应用性很强的学科，在教学设计中要充分考虑统计的学科特点，高中学生的心理特点，认知水平，激发学生学习统计的热情．有条件的可组织学生进行一次具体的抽样调查，从设计调查方案到抽样调查的实施，均由学生独立完成，以培养学生学习的热忱以及分析和解决实际问题的能力．

3．要注意总体、个体和样本的联系和区别．总体均值、方差与样本均值、方差的区别和联系．样本的均值与方差是随机的，我们可以通过合理的抽样，用样本的均值与方差来估计总体的均值与方差．均值与方差是统计学中很重要的两个概念，一个反映总体或者样本的平均程度，另一个反映总体或者样本的离散程度．

4．随机抽样的必要性与合理性．从现实生活中把握必要性，从统计思想中体会合理性．引导学生认识正确设计抽样方案的重要性．

5．在统计学习中，看图是一项十分重要的工作．要引导学生能正确读懂一些统计图表，比如说频率分布图，频率分布直方图，数据散点图等等，从中吸取有用的信息．要求学生能绘制一些简单的统计图表．

6．要注意本章与以前各章的联系与区别，即统计与数学的联系与区别．统计是一门方法论的学科，注重分析和解决实际问题，但一个好的方法论必须有它的理论背景，而数学正是它的一个基石之一．所以在教学中既不能将他们混为一谈，又不能完全割裂它们．

五、评价建议

统计是一门实践性很强的学科，当前世界统计主流以从统计理论研究为主转为统计应用为主．积极参与社会实践活动，设计抽样调查方案，进行抽样调查．根据得来的样本去分析总体的性质，分析整理数据，分析数据间的相关关系，应该作为为评价的重点．例如：

调查问卷部分：教师给定题目后，由学生独立或者分组完成各自的调查问卷的设计，并在一个许可的范围内进行问卷调查．问卷调查全部完成以后，教师首先可以组织学生参照问卷调查的原则对各个或者各组的问卷调查方案进行互评，然后再通过问卷调查的实际效果进行评定．这样既激发了学生的学习兴趣与热忱，又锻炼和提高了学生们运用统计知识解决实际问题的能力．

数据相关性部分：可让学生分析数学成绩与物理成绩之间的相关性，以及语文成绩与数学成绩之间的相关性等等，然后得出结论，比如说数学成绩好的同学物理成绩也不差，而语文成绩与数学成绩之间的相关性没有那么强．让学生分析处理数据，写一篇论文．

教师可以根据实际情况选出学生感兴趣的课题让学生尝试研究．比如，假日经济，旅游经济和休闲时间的统计研究，新药的临床实验研究等等．

学生习题完成情况与测验成绩也可以作为评价的一个方面．

12.1 总体和个体

12.1.1 总体、个体和总体均值

教材线索

本小节从日常生活中涉及总体和个体的实例入手，引出总体、个体、样本和总体均值等基本概念. 最后通过练习和习题引出均值的两个简单性质.

教学目标

1. 理解总体、个体的含义.
2. 理解总体均值的含义.
3. 理解均值的简单性质.

教材分析

1. 重点

总体、个体和总体均值的概念.

2. 难点

(1) 在实际问题中把握总体和个体的含义，问题的侧重点不同，总体和个体可能也不同.

(2) 均值的一些简单性质.

3. 对总体和个体概念的把握

所研究对象的全体元素组成的集合，称为总体（population）. 总体中每一个元素称为个体（individual）.

总体和个体是和研究目的相关的. 比如要调查全校学生的期中测试成绩，则全校学生的期中测试成绩为总体，每一个学生的成绩为个体. 若只关心语文成绩，此时，研究的目的发生了变化，相应的总体和个体也发生变化. 此时，全校学生的语文成绩为总体，而每个学生的语文成绩为个体. 下面再以几个例题说明之.

例 1. 有一批产品共 1 000 个，每个产品可以区分为一等、二等、次品. 我们要研究这批产品的质量，1 000 个产品的等级构成一个总体，每个产品的等级是个体.

例 2. 为考察在某种工艺条件下织出的一批布匹的瑕点数，共取 5 000 匹布. 那么这 5 000 匹布的瑕点数的全体构成一个总体，每匹布的各自瑕点数是个体.

例 3. 在检查某军工厂生产的一大批炮弹的质量时，若只考察炮弹的射程，那么，这批炮弹中每一颗的射程的全体构成一个总体，每颗炮弹的各自射程是个体.

从例 2、例 3 可见，总体中的元素常常不是元素本身，而是指元素的某种数量指标. 在例 2 中，总体中元素指每匹布的瑕点数，在例 3 中，总体中元素指每颗炮弹的射程. 在例 1 中，如果一等品用“1”表示，二等品用“2”表示，次品用“0”表示，同样，总体可看成数“1”，“2”，“0”的集合.

总体的个数可能为有限，也可能为无限. 相应的总体称为有限总体和无限总体.

需要指出的是，在统计学中，人们关心的是研究对象的某项数量指标，这些数量指标在测试前虽然能预料其可能的结果，但确切值是未知的，它们是一个随机变量. 对总体的研究，主要归结为对总体（即某个随机变量）的分布函数及其数量特征的研究.

4. 对总体均值的理解

均值在统计学中是一个十分重要的概念. 总体均值反映了总体数量指标的平均程度.

教学建议

1. 在本小节的教学中，要通过多个实例来让学生把握总体和个体的含义.

2. 引导学生发散思维. 在教材习题 1 中，若每行的数据个数不相等，是否还有 $\mu=(\mu_1+\mu_2+\mu_3)/3$？一般情况又如何？

12.1.2 样本与样本均值

教材线索

本小节首先通过一个尝菜的小例子，初步阐释用样本推断总体的思想．然后介绍样本、样本量、抽样以及样本均值等基本概念．最后通过一个问题，为后面的怎样合理进行抽样调查埋下伏笔．

教学目标

通过具体实例，了解样本（观测数据）、样本量、抽样的概念．理解样本均值的含义．理解只要抽样合理（随机抽样），样本量够大（统计学中的大样本理论），样本均值就是总体均值的一个好的估计．初步体会估计所蕴含的统计思想．

教材分析

1. 重点

样本与样本均值的理解，估计的含义．

2. 难点

样本均值作为总体均值的合理性；怎样选择样本才能使估计更精确．样本均值的随机性的特点．

3. 样本、样本量与抽样

样本，又叫子样，是总体中抽出的一部分个体的集合．样本的个体数量叫样本量．取得这些样本的过程称之为抽样．

4. 总体与样本的关系

总体可以为样本，称之为全样本．但样本一般不是总体，而是总体的一个子集．

5. 样本均值

样本均值是样本的平均值．

假定我们从总体中抽取 n 个个体，记为 x_1，x_2，…，x_n，则（x_1，x_2，…，x_n）称为样本，n 为样本量，记 $\bar{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$，则 $\bar{x}$（读作 x bar）为样本均值. 可以向学生介绍求和符号 $\sum$，音 sigma，可以使式子变得简单.

6. 样本均值与总体均值的把握，估计的含义的理解

（1）总体均值一般是固定的，且是未知的.

（2）样本均值依赖于样本的选取，不同的样本有不同的样本均值，样本均值具有随机性．但是选取样本以后，样本均值可以通过计算得出，因而样本均值是已知的．

（3）样本是从总体中抽出的，总体的特性被样本所保留，因而样本在某种程度上可以代替总体．只要抽样合理，样本量较大，样本均值是总体均值的近似，此时我们称样本均值为总体均值的估计．这就是从已知推断未知，统计学的中心任务．

（4）要注意把统计和精确的数学加以区别，样本均值只是总体均值的一个近似，也就是说样本均值只是总体均值的一个估计，它和总体均值是有一定的偏差的．我们不可能消除偏差，而是只能采取合理的方法对它进行控制．因此，教材中通过一个问题引出一个有偏的估计，为后面抽样调查的重要性埋下伏笔．

（5）教材中讲到一个好的估计所要具备的条件：一是要抽样合理，二是要求样本量较大．我们可以这样理解：所谓“抽样合理”，实际上是要求等可能抽样，简单的说就是要均匀抽样，否则可能导致有偏的结论．关于样本量的问题，若样本量太小，此时我们得到的样本均值随机性太强，不宜作为总体均值的估计．若我们能取到总体中所有的个体，即全样本，此时的样本均值就是总体均值，但是在实际情况中我们不可能也不必取到整个总体．为了得到一个较好的估计，我们只能增大样本量，至于多大才合适，则可根据实际问题而定，也即在实际情况允许的情况下，样本量尽可能的大．增大样本量是一个消除样本均值随机性的好办法．

7. 中位数

我们知道均值是反映总体或者样本平均的一个指标．除此之外，还有一个指标与平均相联系，那就是中位数．具体定义为：

将样本 x_1，x_2，…，x_n 按从小到大（或者按从大到小）排列后为 $x_{(1)}$，$x_{(2)}$，…，$x_{(n)}$，则

$$\text{样本中位数}=\begin{cases} x_{(\frac{n+1}{2})}, & n\text{ 为奇数}, \\ \left[x_{(\frac{n}{2})}+x_{(\frac{n}{2}+1)}\right]/2, & n\text{ 为偶数}. \end{cases}$$

当样本较整齐，则用均值反映样本的平均较好．但是当样本受到较严重的“污染”（即样本的个别数据与其他数据相差太远，统计学上又称之为“离群点”），此时用中位数反映平均较好．兹举一例加以说明，教师在教学中加以掌握．

为考察两个班学生对数学知识掌握的平均程度，特举行一次考试．甲班有三名同学缺考，这三名同学的成绩为零分．乙班学生全部参考．若以均值作为主要依据，则我们得到的结论可能有偏差．此时，我们宜把中位数作为一个重要的依据，来进行判断．此例说明，我们在实际中要根据具体问题加以具体分析．比如说，为了去掉污染数据的影响，有时候我们采用去掉一个最大值，去掉一个最小值再求平均的方法．在各种比赛中，评定选手的成绩时常采用这个方法．

教学建议

1. 教师可以通过一些简单实例来引入一些基本概念比如样本与样本均值的教学．

2. 教师要着重讲解估计的含义，以及估计所蕴含的统计思想．

3. 学生可能对统计和确定性的数学之间的区别不太理解．要让学生知道：统计没有最好，只有更好．没有亘古不变的统计方法，也不可能有包罗一切的统计模型．任何统计方法与统计模型都只是对客观世界的一种近似．统计不可能杜绝错误，只是把错误发生的可能性控制在某个可以接受的范围内．这些就是不确定性的统计与确定性的数学的最大的区别.

相关链接

一、各种平均数：让我们数数有几种

在教材中已经学过均值为一个反映总体或者样本平均程度或者集中趋势的量，现在我们再来看一看还有那些量可以反映这种集中趋势.

1. 众数：一个变量或者样本的众数是指出现最多的数的值.

例如：有一个样本为 {1，1，6，4，5，7，5，2，3，5，0}，显然在这组数据中，“5”出现了 3 次，为次数最多的数，因此，这组数据的众数为“5”.

注意：众数不一定唯一！

众数告诉我们，这个值的出现的次数比其它的值出现的次数多．但它并没有告诉我们它较别的数值多的程度．一个由 100 人组成的群体，无论它有 51 个男人（和 49 个女人）或者 99 个男人（和 1 个女人)，其性别变量的众数都是男人．这两种情况是非常不同的，但是众数并不能区分它们．因此，众数掩盖的信息时常比它揭示的要多.

2. 中位数：数到中间的那一个

一个变量或者样本的中位数是这样一个数字，它把观测值分成同等数目的两部分，一半观察值小于等于这个数，而另一半大于等于这个数.

可以通过把观察值从小到大排序，并取中间的数据值就可以找到中位数．但是要注意，当观察值个数为偶数时，中位数为正中间两个数的平均值.

中位数的优点：中位数很好地代表了一组观测值的中点，中位数只需要很少量的计算．中位数对极端值不敏感，在某些情况下这也是一个优点.

中位数的缺点：除了中间值，中位数并未利用其他观测值．这样，它就没有利用数据中的所有信息.

3. 均值：平衡跷跷板

均值的优点：均值的优点是它对每一个观测值都加以利用．这就意味着比起众数与中位数来，它会获得更多的信息.

均值的缺点：比起众数和中位数，计算比较麻烦；均值对极端值敏感．因此如果由于观测误差而产生的极端值会使结果不好.

例如计算班级平均成绩时，若有几个学生缺考，此时的均值并不能完全反映该班的平均成

绩．再比如，一般大赛中计算选手的成绩都要去掉一个最高分和一个最低分，都是为了排除极端值的干扰．

——节选自吴喜之等译：《统计学基本概念和方法》

二、数理统计学简史

数理统计学是研究怎样去有效地收集、整理和分析带有随机性的数据，以对所考察的问题做出推断或预测，直至为采取一定的决策和行动提供依据和建议．它是 20 世纪发展起来的数学分支，但一些概念及方法可上溯至 18～19 世纪．

1763 年，英国贝叶斯《机会学说问题试解》（Essay Towards Solving a Problem in the Doctrine of Chances）出版，其中建立的条件概率贝叶斯定理成为统计推断的理论基础．1805 年勒让德与高斯同时分别发现最小二乘法，成为统计观测数据误差分析的重要方法．高斯引进的正态分布长期是统计学的基本思想．英国高尔顿（F. Galton，1822—1911）又引入相关与回归概论．

英国数学家皮尔逊（K. Pearson，1857—1936）是现代统计学的奠基人之一．他将统计学用于生物遗传和进化诸问题，得到生物统计学和社会科学统计的一些基本法则．例如，引进了以他名字命名的分布族，包含了正态分布及现在已知的一些重要的偏态分布，还引进矩估方法，用来估计他引进的分布族中的参数，一直是一种重要的参数估计方法．1900 年，他主持创办了《生物统计学》（*Biometrika*）杂志，至今仍是统计学的重要期刊之一．同一年，他提出了检验拟合优度的 χ^2 统计量，并证明其极限分布（在原假设成立时）是 χ^2 分布．这一结果成为大样本统计的先驱工作．皮尔逊的学生戈塞特（W. S. Gossett，1876—1937）引入均值、方差、方差分析、样本等基本概念，还于 1908 年提出 t 检验与 t 分布，开了小样本理论研究的先河．

英国另一数学家歇尔（R. A. Fisher，1890—1962）对现代数理统计的形成和发展做出很大贡献．他系统发展了正态总体下种种统计量的抽样分布（20 年代），标志着相关、回归分析和多元分布等分支的初步建立；建立了以最大似然估计为中心的点估计理论（1912—1925）；合作创立了试验设计，并发展了与这种设计相适用的数据分析方法——方差分析法．他还在假设检验与多元统计分析中做出了创见．

20 世纪 40 年代后，由于经济和军事技术的快速发展，以及电子计算机的出现，使数理统计学的应用达到了前所却未有的规模．例如，罗马尼亚－美国数学家瓦尔德（A. Wald，1902—1950）在《统计决策函数》（Statistical Decision Functions，1950）中创立了统计决策理论，对战后数理统计各分支产生不同程度的影响．

——摘自王青建：《数学史简编》

12.1.3 方差和标准差

教材线索

教材通过一个拔河比赛的具体例子来引入方差的概念．然后介绍总体方差与样本方差的定义以及方差的计算公式，最后介绍标准差的概念．并通过一些例题阐述如何通过方差比较两组样本数据．

教学目标

1．理解总体方差的含义．

2．通过实例理解样本数据方差与标准差的意义与作用．

3．会计算样本数据的方差与标准差．

4．理解样本方差的性质．

教材分析

1．重点

（1）样本方差与标准差的含义与作用．

（2）样本方差与标准差的计算．

2．难点

（1）方差所体现的统计含义，根据定义理解为什么方差是反映总体或者样本中个体整齐程度的一个量，定义的合理性．

（2）方差与标准差的计算．

（3）方差的性质．

3. 我们已经学过均值的概念，知道这是一个反映平均值的量，在很多场合，我们知道这个平均值就可以了．但是均值只反映了数据的平均值，有很大的局限性．在某些场合，仅仅知道均值是不够的．教材通过一个具体实例体现了这一点．虽然两个队的平均年龄是一样的，但是我们无法仅依据这个量来判别这两个队的整体实力．这里还有一个反映数据整齐程度的量，这就是方差．我们再看一例：

例 1. 射手甲和射手乙各射击 10 次，得到成绩如下表所示：

甲	9	10	1	6	4	8	9	0	9	7
乙	5	7	8	8	6	4	8	5	6	6

从上表我们可以计算出，射手甲和射手乙的平均成绩都是 6.3 环. 但是我们是否就可以据此认定这两个射手的水平差不多呢？显然，我们可以看出，虽然这两个射手的均值一样，但是射手乙发挥稳定得多，射手甲甚至还有脱靶的现象. 所以我们还需要一个反映数据波动程度或者整齐程度的量，即方差.

4. 总体方差的理解

（1）定义：记 y_1，y_2，…，y_N 是总体中的全部个体，μ 为总体均值，则总体方差定义为

$$\sigma^2=\frac{1}{N}\sum_{i=1}^{N}(y_i-\mu)^2.$$

可见，总体方差是总体的平均平方误差. 加上平方是为了避免和式中正负相抵的情况出现，方差中的“方”就是指平方，“差”是指每个个体与均值的差. 方差越大，总体中个体偏离总体均值的程度就越大；方差越小，这种偏离的程度就越小，数据就越整齐，当方差为零时，总体中的每个个体都为同一值了.

（2）是否还有其他的量也可以反映这种数据的整齐程度呢？答案是肯定的. 把方差定义中的平方改成绝对值或者其他任意偶数次方，都可以反映这种数据的波动程度. 我们采用平方主要是为了数学上处理的方便.

（3）总体方差的未知性.

在方差定义中，由于总体均值是未知的，因而总体方差也是未知的，需要估计. 下面介绍的样本方差就是总体方差的估计.

5. 样本方差的理解

（1）定义：给定 n 个观测数据 x_1，x_2，…，x_n，用 $\bar{x}$ 表示样本均值，则样本方差定义为：

$$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

与总体方差一样，样本方差也反映了样本数据的整齐程度，是描述数据的发散程度或者波动程度的指标.

（2）样本方差的随机性与已知性.

显然，样本方差与样本均值一样，依赖于样本的选取，不同的样本具有不同的样本方差，因而具有随机性. 样本方差可以由样本计算出，因此是已知的，正因为如此，我们可以用样本方差来估计总体方差. 注意，样本方差的定义中不能用总体均值 μ 代替样本均值 $\bar{x}$，否则的话，样本方差也成未知的了.

6. 多知道一点：样本方差

限于学生的理解和认知水平，教材中定义的样本方差实际上是总体方差的一个有偏估计，即 $E(s^2)=\frac{n-1}{n}\sigma^2\neq\sigma^2$，“$E$”表示数学期望，也即平均的意思. 在统计学中，真正的样本方差是如下定义的：

$$s_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2,$$

这时，$E(s_1^2)=\frac{n}{n-1}\cdot E(s^2)=\frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2=\sigma^2$，即此时的样本方差才是总体方差的无偏估计．在统计学中，估计的无偏性是一个基本的要求．当然，在样本量 n 很大的情况下，这两个样本方差的差别是很小的，当 $n\to+\infty$时，我们有 $E(s^2)=\frac{n-1}{n}\sigma^2\to\sigma^2$，即 s^2 是总体方差 σ^2 的渐近无偏估计．

7. 方差的计算公式

按教材方差的定义可以计算出方差，但是比较繁琐，因为每一个数据都要做中心化处理，即都要减去样本均值．教材中给出了方差计算的另一个公式，即：

$$s^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2,$$

这两个式子是等价的，后一个式子计算起来简便一些，学生应该牢固掌握．在一些带有统计功能的计算器中，只要输入数据，就可以立即算出样本方差．在学生使用带统计功能的计算器计算样本方差时，实际上是按照我们后面给出的无偏估计的定义算的，算出来的是 s_1^2 而非 s^2，所以算出来的结果是和学生按教材上的方法算出来的结果是有差异的，教师应该明确这一点．当然，在一些常用的统计软件如 S－plus，Matlab，SAS 中也是按 s_1^2 计算样本方差的．

例 2. （接例 1）计算射手甲和射手乙的方差，对两个射手进行评价．根据定义，我们可以算出射手甲的方差为 11.21，射手乙的方差为 1.81．显然，射手乙发挥更稳定．注：按照无偏估计计算出射手甲和射手乙的方差分别为 12.455 56 和 2.011 111.

8. 标准差

标准差是为了统一量纲时引入的，它是方差的算术平方根，也是描述数据发散程度或者波动幅度的指标．

9. 方差的性质

在前面学习均值的性质时，我们知道若每个数据都增加相同的值，则样本均值也增加相同的值．对于样本方差，这个性质不再有了，即当每个数据都增加相同的值时，样本方差不变！教材中练习明确了这一点．事实上，每个数据都增加相同的值，数据的波动程度并没有发生改变！

但是，当数据都放大 a 倍时，样本均值依然放大 a 倍，而方差却放大 a^2 倍，样本标准差放大 $|a|$ 倍，学生应该注意到这些区别．

教学建议

1. 建议教师通过丰富的实例来引入方差的教学，增加学生对方差的直观了解．

2. 方差的定义没有均值那么直观，是教学的难点之一．教师要重点讲清楚定义的合理性，使学生能够理解．

3. 要通过多做练习使学生掌握方差的计算．有条件的话，可以让学生通过带有统计功能的计算器或计算机进行计算．

4. 教师要结合均值的内容让学生理解为什么样本均值可以作为总体均值的估计以及要作为一个好的估计必须具备哪些条件.

5. 教材习题 3 给出了一个例子，说明在产品制造业中，人们往往更关心产品的方差. 对工厂产品抽样检测，抽检产品的方差是一个重要指标. 教师在讲授时，要多联系实际，让学生知道统计是一门非常有用的学科，在我们生活中无所不在，激发学生学习统计的兴趣和下定学好统计的决心.

例题分析

例 1. 研究某地 10 岁男孩的身高的分布情况，抽取了 60 名男孩进行测量，得到如下数据（单位：cm)，请计算这组样本的均值，方差与标准差.

128.1	144.4	150.3	146.2	140.6	139.7
134.1	124.3	147.9	143.0	143.1	142.7
126.0	125.6	127.7	154.4	142.7	141.2
133.4	131.0	126.4	130.3	146.3	146.8
142.7	137.6	136.9	122.7	131.8	147.7
135.8	134.8	139.1	139.0	132.3	134.7
138.4	136.6	136.2	141.6	141.0	138.4
145.1	141.1	139.9	140.6	140.2	131.0
150.4	142.7	144.3	136.4	134.5	132.3
152.7	148.1	139.9	138.6	136.1	135.8

解： 用 μ，s^2，s 分别表示样本均值，样本方差与标准差.

$$\mu = \bar{X} = \frac{1}{60}\sum_{i=1}^{60} X_i \approx 138.55,$$

$$s^2 = \frac{1}{60}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = 19247.59 - 138.55^2 \approx 51.49,$$

$$s = \sqrt{s^2} = \sqrt{51.49} \approx 7.18.$$

相关链接

数理统计学的形成

统计的历史可以说和人类的历史一样长. 人口、财富、军事等的统计是和父系氏族社会的需要相联系的. 没有一个国家可以不需要统计. 但是，使用数学方法研究统计规律，则是近代的事.

英国地质学家莱尔（Charles Lyell，1797－1875）于1830－1833年间出版了三卷《地质学原理》，他根据各个地层中的化石种类和现仍在海洋中生活的种类做出百分比，然后定出更新世、上新世、中新世、始新世的分期．这些名称沿用至今，莱尔称为地质学上一位划时代的人物．然而，他所用的方法——类似于数理统计方法——却被人遗忘了．1831年，达尔文（C. R. Darwin，1809－1882）从剑桥大学神学院毕业后，参加英国海军测量船“猎犬号”作环球实地考察．他在船上带有莱尔的《地质学原理》．达尔文在收集、整理大量标本的过程中，借助了生物计量方法，即统计方法．人们只关心达尔文进化论的结果，也忽视了他所用的方法．其实，达尔文的工作的确是数理统计学的先声．生物统计学正是现代数理统计学的摇篮时期，它的故乡，也正是英国．

生物统计学派的奠基人是英国的高尔顿（Francis Galdon，1822－1911）．他毕业于剑桥的三一学院．表哥达尔文的巨著《物种起源》问世以后，触动他用统计方法研究智力遗传进化问题，第一次将概率统计原理等数学方法用于生物科学，明确提出“生物统计学”的名词．现在大家耳熟能详的“相关”和“回归”，也是高尔顿第一次使用的．1870年，高尔顿在研究人类身高的遗传时，发现下列的关系：高个子父母的子女，其身高有低于其父母身高的趋势，而矮个子父母的子女，其身高有高于其父母的趋势，即有“回归”到平均数去的趋势．这就是统计学上最初出现“回归”时的涵义．

真正从数学上开始生物统计研究的第一人，则是皮尔逊（Karl Pearson，1857－1936）．他在伦敦英王学院攻读数学．1884年任伦敦大学的应用数学与力学教授．1890－1900年间，皮尔逊在高尔顿的指点下，讨论生物进化、返祖、遗传、自然选择、随机交配等问题，用回归和相关工具，系统地将生物进化数量化，并先后提出和发展了标准差、正态曲线、平均变差、均方误差等一系列数理统计学的名词和概念．1901年，皮尔逊创办《生物计量学》（*Biometrica*）杂志，这是数理统计学的第一份期刊，现在已是统计学中的“四大天王”杂志之一．这位生物统计学的开创者，对1920年以后的新事物缺乏敏锐的感觉，和“小样本”理论发生激烈的冲突，直至去世．

皮尔逊致力于进化论，数据少则几十，多则数百．为了结论的正确，样本总是越大越好．但是在现实世界中，并非都能获得大样本．如果样本量很小，统计工作如何进行？这副担子落在了一位酿酒化学技师的身上．

1899年，在牛津大学学习数学和化学的戈塞特（W. S. Gossett，1876－1937），来到都柏林市的一家酿酒公司担任酿酒化学技师，从事统计和实验工作．酿酒公司请一位有数学方面的专长的人来当技师，表明当时英国社会和企业界对统计的重要性已有相当认识．公司里，酿酒的技师当然不少，戈赛特就想用自己擅长的数学，从统计上搞出一点成绩来．他首先注意到，供酿酒的每批麦子的质量相差很大，同一批麦子中能抽样供实验的麦子又很少，每批样本在不同的温度下做实验，其结果相差很大．这样一来，实际上取得的样本，只能是小样本．可是，从小样本来分析数据是否可靠？误差有多大？小样本理论就在这样的背景下孕育着．

1907年，戈赛特决心把小样本和大样本之间的差别搞清楚．为此，他试图把一个总体中所有

小样本的平均数的分布刻画出来．这个分布就是 t 分布函数．1908 年，戈赛特以“*Student*（学生）”的笔名在《生物计量学》上发表了论文《平均数的随机误差》．这篇文章提供了“学生” t 检验的基础，开创了统计学的小样本理论，被统计学家看作统计推断理论发展史上的里程碑．

戈赛特在酒厂里干了一辈子，但真正留给世人的财富却是划时代的统计工作．一般认为，戈赛特的研究工作的意义不仅在于获得小样本理论，更重要的是他的研究方法和思考路线，即不仅是依靠大量观察进行描述，而是通过实验思考得出推断．

戈赛特的小样本理论发表之后，一时未获承认．当时公认的分布是正态分布，代表当时统计学最高水平的伦敦大学生物统计研究所，只研究有关大样本的各种问题．当时的统计权威皮尔逊对小样本理论持有怀疑态度．他觉得统计学的基本精神是要求大样本，小样本的倾向是危险的．这样，戈赛特的 t 分布理论只有在他所服务的酿酒公司里用于实验分析和业务管理，并没有为外界所接受．

下一个开创数理统计学新局面的是费歇尔（R. A. Fisher，1890－1962）．费歇尔出生在一个没落的拍卖商家庭．1909 年，靠一笔助学金进入剑桥大学附属的一个学院读书．对他影响最大的是皮尔逊，《生物计量学》是他最喜爱的刊物．1919 年，皮尔逊推荐费歇尔到罗萨姆斯太德农业站任职，第一线的工作给费歇尔带来了丰富的实验数据和资料．他的工作不仅推动了农艺工程的进步，更促进了统计学的发展．在二十世纪二十至四十年代，罗萨姆斯太德实验站成了世界“推断统计”的中心．戈赛特的推断统计在这里得到了充分重视．

费歇尔发现戈赛特的 t 分布在分析实验结果时非常有用．1923 年，费歇尔又对 t 分布作了严密而简单的理论证明．到了 1925 年，费歇尔在他的名著《供研究人员用的统计方法》一书中，附有“学生” t 分布概率表，戈赛特的工作获得了迅速的传播和发展，其学术价值得到了举世公认．

1933 年，费歇尔在罗萨姆斯太德实验站工作 14 年之后，回到了剑桥大学任优生学系和遗传学系教授．1959 年，69 岁高龄的费歇尔退休，此后来到澳大利亚的联邦科学与工业研究组织中担任了一些统计工作，度过了一生的最后三年．

费歇尔一生没有担任过统计学教授，却是他那个时代的统计学领袖．在他手里，数理统计学脱离生物计量学的范围获得独立．费歇尔的工作几乎涉及到数理统计学的一切领域．他所提出的 Z 分布由他的学生改进被称为 F 分布（用费歇尔的第一个字母命名）是统计学中三大分布之一．现在广泛使用的方差分析、实验设计、参数估计，都是他首创的．

英国的数理统计学派在二十世纪上半叶一直处于领先地位．世界各国的统计学家，包括后来著名的美国统计学家奈曼（J. Neyman，1894－1981）都来到伦敦大学学院．我国著名的概率统计先驱许宝禄（1910－1970）于 1936 年来到伦敦大学，在那里获博士学位．他在多元分析、极限定理方面有杰出的工作．对奈曼－皮尔逊理论的建立和发展起了重要作用．

1940 年以后，以奈曼来到美国为标志，英国数理统计学派开始衰落，美国逐渐成为新的数理统计学的研究中心．

——节选自张奠宙：《20 世纪数学经纬》

12.2 抽样调查方法

12.2.1 随机抽样

教材线索

本节教材通过几个简单有趣的例子说明我们在日常生活中经常用到抽样调查的方法，再通过一例引出抽样调查的必要性. 然后通过回答 12.1.2 节的问题，引出如何设计抽样调查方案，以及抽样调查的概念. 最后介绍简单随机抽样与简单随机样本和随机数的概念.

教学目标

1. 了解抽样调查的必要性与必须性.
2. 了解抽样调查的含义，知道怎样合理的设计抽样调查方案.
3. 了解有放回随机抽样与无放回随机抽样的区别，知道在实际生活中选取抽样方法.
4. 了解随机数的含义，会根据随机数表进行随机抽样.

教材分析

1. 重点

合理的设计随机抽样方案进行随机抽样.

2. 难点

（1）随机抽样方案的设计；

（2）随机抽样思想的领悟.

3. 抽样调查的可能性

教材通过一个品尝汤的味道的例子来说明通过抽样调查取得样本，并用样本来推断总体是可能的. 并说明好的抽样调查应具有的两个条件：

（1）随机性. 这是抽样调查成功的关键！

（2）合适的样本量. 样本量太少不足以反映总体的情况，样本量太大又可能受到客观条件的限制或者实际问题不需要.

这两个条件是相辅相成的，缺一不可.

抽样调查的可能性是从统计学的层面上来说的.

4. 抽样调查的必要性

（1）在实际问题中，我们要推断总体的性质，但我们往往很难取到全部总体，或者取到全部总体所耗费的人力物力太大. 比如说要了解某市居民的健康状况，收入水平等，不可能对所有市民都进行调查.

（2）有些试验，本来就是带有破坏性的. 比如要了解一批炮弹的射程，如果我们把全部炮弹都试射了，虽然我们得到了精确的数据，但这一批炮弹也都报废了. 同样的例子还有机器或者灯管的使用寿命等.

（3）有些试验，本身就是非常昂贵的，比如新药的临床试验等.

这些都说明，抽样调查不但是必要的，而且也是必须的.

抽样调查的必要性是从实际应用的层面上来说的.

5. 如何合理的设计抽样调查的方案呢？关键是要随机抽样！

所谓随机，就是要使总体中的个体都等可能的被抽到，即每个个体被抽到的机会都相同，否则可能会导致有偏的估计. 比如说教材 12.1.2 中的例子，女生倾向于调查女生，从而导致有偏的结论. 抽样时不能带有任何主观的成分，最好是采用随机数表或者计算机产生随机数的方法进行抽样.

教材通过阅读材料“《文学摘要》的破产”阐释了随机抽样的重要性. 有偏抽样最终可能导致错误的结论. 所以我们在设计抽样方案时一定要注意抽样的随机性.

6. 关于简单随机抽样

简单随机抽样是指不放回随机抽样.

（1）它具有广泛的实用性，抽样实践中多采用不放回抽样，而且抽取的样本中没有重复的抽取的个体，便于进行有关的分析和计算.

（2）它是从总体中逐个地进行抽取，有利于在实践中进行操作，比如说抽签法就是一种常见的简单随机抽样.

（3）它是一种等概率抽样. 在整个抽样过程中，每个个体被抽取的概率相等，从而保证了抽样的公平性. 比如说总体有 N 个个体，从中抽取一个容量为 n 的样本时，采用无放回随机抽样，则每个个体被抽到的概率相等，都等于 n/N. 它是具有严格的概率意义的，由于要说清楚这件事情需要用到条件概率，因此我们在此不证明这个事实.

7. 关于“增大样本量可以提高估计的精确度”的理解

从直观的层面上来看，样本量越大，样本所包含总体的信息越多，因而估计越精确.

从统计意义上来说，估计的精确度是与估计的方差相联系的. 一个估计比另一个估计的精确度更高，是指这个估计的方差比另一个估计的方差更小. 设样本均值 $\bar{x}$ 是总体均值 μ 的估计，我们可以计算出估计的方差 $\mathrm{var}(\bar{x})=\mathrm{var}(\frac{1}{n}\sum_{i=1}^{n}x_i)=\frac{1}{n^2}\mathrm{var}(\sum_{i=1}^{n}x_i)=\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{var}(x_i)=\frac{1}{n}\sigma^2$，计算中我们假定 x_i 是独立同分布的，σ^2 为其方差. 可见，当样本量 n 增大时，估计的方差就变小，从而估计就更精确.

教学建议

1. 建议教师多通过几个实例阐述抽样调查的必要性.

2. 随机抽样重在“随机”两字上，教师要把这点讲解清楚. 教师可以通过教材 12.1.2 的例子和阅读材料来说明非随机或者随机性不强的不利后果.

3. 简单随机抽样是一种无放回随机抽样，实践中最简单的一种方法就是抽签法. 教师可以在课堂上做一些试验.

相关链接

抽样框

在正式进行抽样调查前，当总体中个体数目很多时，往往必须事先编制好抽样框. 下面我们介绍什么是抽样框.

为了抽样的方便往往把总体中的个体进行归并，一个个体或者若干个个体组成了一定的个体集合. 我们称这些个体所组成的集合为抽样单元. 而总体就是由这些抽样单元组成. 抽样单元可大可小，大的抽样单元可以由小的抽样单元组成. 也就是说，抽样单元可以分成级. 对总体进行第一次划分得到初级单元或第一级单元，对初级单元再进行划分就得到次级单元或者第二级单元. 同样，可以定义三级单元、四级单元等等.

把所有基本单元放在一起编上号码就构成了总的抽样框. 当然，在一个上级单元中给它的所有下级单元编上号码也可以得到它自己的抽样框. 抽样框相当于一份名册或清单. 它的作用主要是按一定随机化程序进行对下级单元的抽样. 随机化程序抽样可以通过查随机数表，或掷随机数骰子，或在计算机上产生“随机数”等办法进行. 一般地，对于人口普查和人口抽样调查，由于政府部门出面会有令人满意的行政和组织措施，拥有现成的抽样框. 但是，对于有些调查来说，比如市场调查涉及的商品的花色、品种、规格实在太多，有时候难有现成的抽样框，因此需要在实施抽样调查前另行编制抽样框.

为了确保样本的随机性，抽样框中的单元一旦入选样本，调查员无权自作主张更改!

12.2.2 调查问卷的设计

教材线索

教材首先通过一个具体例子（敏感问题调查）来说明如何合理地设计调查问卷，然后再说

明正确的调查的方式所要注意的几个问题.

教学目标

1. 通过实例，了解调查问卷的设计要求.
2. 能在实践中设计一些简单的调查问卷.

教材分析

1. 重点

调查问卷的设计.

2. 难点

合理的调查问卷的设计，怎样设计调查问卷才能收到预期的效果.

3. 在调查问卷设计中，合理地设计调查问卷方案是非常重要的，一个好的调查方案应该能很好地达到调查的目的，因此，设计时应该遵循一些原则，比如合理性、一般性、逻辑性、明确性、非诱导性和便于整理分析等，在本节后面的相关链接中详细地阐述了这些原则，供教师在教学中参考.

教学建议

在教材分析里面，我们给出了调查问卷设计的一般原则，教师可以拟定一个与学生学习和生活息息相关的内容，要求学生设计一份调查问卷，并要求学生课外完成.

相关链接

一、调查问卷的设计原则

问卷调查是目前调查业中所广泛采用的调查方式——即由调查机构根据调查目的设计各类调查问卷，然后采取抽样的方式确定调查样本，通过调查员对样本的访问，完成事先设计的调查项目，最后，由统计分析得出调查结果的一种方式. 它严格遵循的是概率与统计原理，因而，调查方式具有较强的科学性，同时也便于操作. 这一方式对调查结果的影响，除了样本选择、调查员素质、统计手段等因素外，问卷设计水平是其中的一个前提性条件. 而问卷设计的好坏很大程度上又与设计制度（原则）有关！这里补充一些调查问卷设计的原则，供教师在教学中参考.

（1）合理性. 合理性指的是问卷必须紧密与调查主题相关. 违背了这样一点，再漂亮或精美的问卷都是无益的. 而所谓问卷体现调查主题其实质是在问卷设计之初要找出与“调查主题相关的要素”！

如："调查某化妆品的用户消费感受"——这里并没有一个现成的选择要素的法则．但从问题出发，特别是结合一定的行业经验与商业知识，要素是能够被寻找出来的．一是使用者（可认定为购买者）．包括她（他）的基本情况（自然状况：如性别、年龄、皮肤性质等）；使用化妆品的情况（是否使用过该化妆品、周期、使用化妆品的日常习惯等）．二是购买力和购买欲．包括她（他）的社会状况收入水平、受教育程度、职业等）；化妆品消费特点（品牌、包装、价位、产品外观等）；使用该化妆品的效果（评价．问题应具有一定的多样性、但又限制在某个范围内，如价格、使用效果、心理满足等．三是产品本身．包括对包装与商标的评价、广告等促销手段的影响力、与市场上同类产品的横向比较……应该说，具有了这样几个要素对于调查主题的结果是有直接帮助的．被访问者也相对容易了解调查员的意图，从而予以配合．

（2）一般性．即问题的设置是否具有普遍意义．

应该说，这是问卷设计的一个基本要求，但我们仍然能够在问卷中发现这类带有一定常识性的错误．这一错误不仅不利于调查成果的整理分折，而且会使调查委托方轻视调查者的水平．如搞一个"居民广告接受度"的调查：

问题：你通常选择哪一种广告媒体？

答案：a. 报纸；b. 电视；c. 杂志；d. 广播；e. 其他

而如果答案是另一种形式

a. 报纸；b. 车票；c. 电视；d. 墙幕广告；e. 汽球；f. 大巴士；g. 广告衫；h. ……

如果我们的统计指标没有那么细（或根本没必要），那我们就犯了一个"特殊性"的错误，从而导致某些问题的回答实际上是对调查无助的！

在一般性的问卷技巧中，需要注意的是：不能犯问题内容上的错误．如：

问题：你拥有哪一种信用卡？

答案：a. 长城卡；b. 牡丹卡；c. 龙卡；d. 维萨卡；e. 金穗卡

——其中"d"的设置是错误的，应该避免．

（3）逻辑性．问卷的设计要有整体感，这种整体感即是问题与问题之间要具有逻辑性，独立的问题本身也不能出现逻辑上的谬误．从而使问卷成为一个相对完善的小系统．如：

问题：

1. 你通常每日读几份报纸？

a. 不读报；b. 1 份；c. 2 份；d. 3 份以上

2. 你通常用多长时间读报？

a. 10 分钟以内；b. 半小时左右；c. 1 小时；d. 1 小时以上

3. 你经常读的是下面哪类（或几类）报纸？

a. ×市晚报；b. ×省日报；c. 人民日报；d. 参考消息；e. 中央广播电视报；f. 足球……

在以上的几个问题中，由于问题设置紧密相关，因而能够获得比较完整的信息．调查对象也会感到问题集中、提问有章法．相反，假如问题是发散的、带有意识流痕迹的，问卷就会给人以随意性而不是严谨性的感觉．那么，将市场调查作为经营决策的一个科学过程的企业就会

对调查失去信心！

因此，逻辑性的要求即是与问卷的条理性、程序性分不开的. 已经看到，在一个综合性的问卷中，调查者将差异较大的问卷分块设置，从而保证了每个“分块”的问题都密切相关.

（4）明确性. 所谓明确性，事实上是问题设置的规范性. 这一原则具体是指：命题是否准确？提问是否清晰明确、便于回答；被访问者是否能够对问题作出明确的回答，等等.

如上文问题中“10 分钟”、“半小时”、“1 小时”等设计即是十分明确的. 统计后会告诉我们：用时极短（浏览）的概率为多少；用时一般（粗阅）的概率为多少；用时较长（详阅）的概率为多少. 反之，答案若设置为“10 分钟～60 分钟”，或“1 小时以内”等，则不仅不明确、难以说明问题，而且令被访问者也很难作答.

再则，问卷中常有“是”或“否”一类的是非式命题. 如

问题：您的婚姻状况

答案：Ⅰ. 已婚；Ⅱ. 未婚；

显而易见，此题还有第三种答案（离婚/丧偶/分居）. 如按照以上方式设置则不可避免地会发生选择上的困难和有效信息的流失！其症结即在于问卷违背了“明确性”的原则.

（5）非诱导性. 不成功的记者经常会在采访中使用诱导性的问题——这种提问方式如果不是刻意地要得出某种结论而甘愿放弃客观性的原则，就是彻头彻尾的职业素质的缺乏. 在问卷调查中，因为有充分的时间作提前准备，这种错误大大地减少了. 但这一原则之所以成为必要，是在于高度竞争的市场对调查业的发展提出了更高的要求.

非诱导性指的是问题要设置在中性位置、不参与提示或主观臆断，完全将被访问者的独立性与客观性摆在问卷操作的限制条件的位置上. 如：

问题：你认为这种化妆品对你的吸引力在哪里？

答案：a. 色泽；b. 气味；c. 使用效果；d. 包装；e. 价格；f. ……

这种设置是客观的. 若换一种答案设置：

a. 迷人的色泽；b. 芳香的气味；c. 满意的效果；d. 精美的包装……

这样一种设置则具有了诱导和提示性，从而在不自觉中掩盖了事物的真实性.

（6）便于整理、分析. 成功的问卷设计除了考虑到紧密结合调查主题与方便信息收集外，还要考虑到调查结果的容易得出和调查结果的说服力. 这就需要考虑到问卷在调查后的整理与分析工作.

首先，这要求调查指标是能够累加和便于累加的；其次，指标的累计与相对数的计算是有意义的；再次，能够通过数据清楚明了地说明所要调查的问题.

只有这样，调查工作才能收到预期的效果.

二、保险市场调查问卷实例

第一部分：甄别部分

（1）请问你的家人有无在保险行业工作的？

a. 有（结束访问）　　　b. 无

（2）请问你的年龄

若年龄在 18 岁～50 岁以外，结束访问

……………………………

第二部分：一般风险意识与投保倾向

Q5. 有的人碰到下列的某些问题，请问你对这些问题是否产生过以下担忧？（在合适处打勾）

	（1）经常	（2）有时	（3）偶尔	（4）不担忧
1. 车祸	________	________	________	________
2. 健康问题	________	________	________	________
3. 退休后的生活保障	________	________	________	________
4. 财产被盗	________	________	________	________
5. 下岗或失业	________	________	________	________

Q6. 下面有些说法，人们的观点可能不同，请你给下面的观点打分：

很同意（5 分）　比较同意（4 分）　讲不清（3 分）　不太同意（2 分）　很不同意（1 分）

1. 越是有钱，越应该参加保险. [　]
2. 年轻人没有必要买养老保险. [　]
3. 只有人们的收入达到一定水平，才会考虑保险. [　]
4. 现在的保险让人信不过. [　]
5. 交费容易、理赔难. [　]
6. 保险推销员让人烦. [　]
7. 保险不如储蓄回报率高. [　]
8. 保险宣传力度不够，人们对具体内容不了解. [　]
9. 目前保险定价比较合理. [　]

Q7.　假如银行利息率为 3%，您认为养老保险的回报率（与银行利息率类似）至少达到____%时您会投保.

Q8. 如果在满足日常消费之后，您另有一笔结余资金 10 万元，根据以下信息，您愿意怎样处理它们？（注意各百分数之和必须为 100%）

投资种类	特征	投资百分比
1. 投资于银行	收益率较低，安全性很高，可以提前支取	
2. 投资于股市	收益率从长期来看比较高，但安全性较差，有可能血本无归	
3. 购买养老保险	中等收益率，安全性较高，但资金一旦投入，养老金就很难支配，如果提前支取就会面临一定损失	
4. 住宅	……	
5. 其他	……	

Q9. 中国加入 WTO，如果外国保险公司也将在您所在地办理保险业务，您愿意选择：

1. 国内保险公司

2. 国外保险公司

3. 中外合资的保险公司

4. 无所谓，根据具体情况而定

Q10. 购买保险您会注重（最多选三个）[　]、[　]、[　].

1. 保险费高低　2. 服务质量　3. 公司信誉　4. 地理位置方便　5. 广告印象　6. 民族感情　7. 亲戚朋友影响　8. 对保险公司的了解　9. 其他（请注明＿＿＿＿＿＿）

Q11. 您认为我国保险公司的前景如何：

1. 好　　2. 一般　　3. 不好

…………………………

——节选自《统计学》教学案例和教学项目汇编

12.2.3 分层抽样与系统抽样

教材线索

本小节通过一个品尝西红柿炒鸡蛋的例子来引入分层抽样方法，然后介绍分层抽样的特点，最后通过一个住宅小区的调查实例介绍系统抽样方法.

教学目标

1. 理解分层抽样方法，知道分层抽样的特点，能实施一些简单的分层抽样.

2. 了解系统抽样方法，会用系统抽样的方法从总体中抽取样本.

3. 通过对本小节的学习，提高学生对统计的认识，进一步提高学生理论联系实际的能力.

教材分析

1. 重点

分层抽样.

2. 难点

(1) 合理的设计分层抽样的方案;

(2) 系统抽样方法的合理性与公平性.

3. 分层抽样的定义

如果总体是由互不重叠的差异明显的几个部分(称每个部分为层)组成,抽样在每一层中独立地进行,总的样本由各层样本组成,根据各层样本汇总对总体进行推断. 这种抽样方法称之为分层抽样方法(Stratified sampling method).

4. 分层抽样的特点和适用场合

分层抽样技术是实际中最常采用的抽样技术之一. 它具有许多其它抽样所没有的特点.

(1) 分层抽样是在各层中进行的,因此各层样本除汇总后可以对总体进行推断外,还可以用来对每层进行推断. 例如一项全国性的抽样调查,若以省为层,那么调查以后既可以得到有关全国的数据,也可同时获得有关省的数据. 又比如要调查高三学生的平均身高,以性别为层,我们既可以得到所有学生的平均身高,又可以分别得到男生和女生的平均身高.

(2) 分层抽样实施起来灵活方便,而且也便于组织,它允许根据不同层的具体情况采用不同的抽样方法.

(3) 因为分层样本分别来自各层,因此与简单随机样本比较,分层样本在总体中的分布更加均匀,不会出现偏于某一部分的不平衡情况. 这也是分层抽样受到实际工作者欢迎的原因之一.

(4) 分层抽样能较大的提高调查的精度.

5. 系统抽样方法

把总体的个体按一定的方式排列,在规定的范围内随机抽取一个个体,然后按照制定好的规则确定其他个体的抽样方法称为系统抽样方法(Systematic sampling method).

从上述定义可见,系统抽样方法的主要优点是实施简单. 与其他随机抽样方法相比,其缺点在于随机性不太强.

6. 系统抽样的实施步骤

(1) 编号(可使用个体本身所带的号码).

(2) 分段(等间隔分段,间隔为 d).

(3) 确定起始个体编号 l(在第一段采用简单随机抽样方法确定).

(4) 按照事先规定的规则抽取样本. 通常第一个为 l,第二个为 $l+d$,第三个为 $l+2d$,依此类推,直至抽样完成. 下面举两个常见的系统抽样的例子.

例 1. 大会报告完了以后要留下一部分人座谈,通知每排的 10 号留下来.

例 2. 工厂抽检产品的质量,在用传送带将产品送入包装车间前,质检人员每隔 10 分钟从

传送带上抽取一件产品进行检验.

7. 简单随机抽样，分层随机抽样与系统抽样的联系与区别

（1）简单随机抽样最基本，其他两种方法都是建立在它的基础之上的；

（2）三种抽样方法都是等概率（可能）抽样，体现了抽样的公平性；

（3）三种抽样方法各有其特点和适用范围，简单随机抽样适用于总体中个体数不多的情况；分层抽样适用于总体由差异明显的几部分组成的情况；系统抽样适用于总体中个体数量较多的情况. 在抽样实践中要根据具体情况选用相应的抽样方法.

教学建议

因在实际中，分层抽样技术用得最广泛，因此，教师应着重介绍分层抽样方法. 可让学生自己举几个在生活中碰到的分层抽样的例子. 由于系统抽样相对简单，在日常生活和生产中也很常用，教师可通过几个实例，让学生判断是属于何种抽样. 也可以让学生举几个系统抽样的例子.

例题分析

例 1. 某校有老师 200 人，男学生 1 200 人，女学生 1 000 人. 现用分层抽样的方法从所有师生中抽取一个容量为 n 的样本；已知从女学生中抽取的人数为 80 人，则 $n=$__________.

解：该校师生总数为 $200+1\ 200+1\ 000=2\ 400$，其中女生所占的层权为$\frac{1\ 000}{2\ 400}=\frac{5}{12}$，现在已知女生已抽取 80 人，则样本量 n 应该为

$$80\div\frac{5}{12}=192\text{（人）}.$$

例 2. 某公司甲、乙、丙、丁四个地区分别有 150 个、120 个、180 个、150 个销售点. 公司为了调查产品的情况，需从这 600 个销售点中抽取一个容量为 100 的样本，记这项调查为①；在丙地区中有 20 个特大型销售点，要从中抽取 7 个调查其收入和售后服务等情况，记这项调查为②. 则完成这两项调查宜采用的抽样方法依次为（　　）

A. 分层抽样法，系统抽样法

B. 分层抽样法，简单随机抽样法

C. 系统抽样法，分层抽样法

D. 简单随机抽样法，分层抽样法

解：为了保证甲、乙、丙、丁四个地区的销售点都被等可能的抽取到，因此，调查①宜采取分层抽样法，而在调查②中，由于总体个数不多，为保证抽样的随机性，宜采用简单随机抽样法. 因此答案选 B.

相关链接

整群抽样

下面介绍一种抽样方法——整群抽样，这也是一种非常重要也是十分常见的抽样方法.

设想湖南教育厅想了解长沙市中学生的体质情况，抽样调查是既省钱又省时的办法. 显然，长沙市的中学生均是总体的单元，从全体学生中随机无放回地抽取若干样本是理想的抽样方法（简单随机抽样），但是，由于总体数量太大，同时一个合理的样本一般要遍布全市，对如此分散的中学生样本逐个进行访问，其工作量之大可想而知. 一个方便的方法是在长沙地区按学校抽样，在抽到的几所学校中对该校所有中学生进行普遍调查. 这就是整群抽样.

若总体科研分为 N 个初级单元（称作群），每个初级单元包含若干次级单元. 按某种方式从总体中抽取 n 个初级单元，对这些单元中的所有次级单元全部进行调查. 这种抽样方法称为整群抽样.

整群抽样的好处在于，倘若想调查的最基本的单元缺乏现成的抽样框，或很难编制抽样框，而这些基本单元的某种组合的抽样框易得，便可以实施. 这种实施手法比较便利，在费用和时间上均有节省.

整群抽样比起简单随机抽样来，样本的代表性要差一些，当然我们可以通过略多抽几个群来弥补这一缺陷. 但是最关键的一条还是在于总体内群的划分. 一般应遵循的原则是：一、群与群之间的差异要尽可能的小（这一点与分层抽样中总体中层的划分有着极大的差别）；二、群内单元之间的差异应相当大，这意味着每个群均具有足够多的代表性. 如果所划分的群相互之间有颇多的相似之处，那么少量群的抽取足以提供良好的精度. 比如说前面的中学生体质调查，每一所中学都可以化为一个群，它们之间互相很相似（关于学生的体质），但在一个学校里每个学生之间有一定的差异. 倘若需要我们自行划分群，一般还要考虑到组织管理上的方便、精度上的要求以及费用多少等因素.

——节选自施锡铨：《调查的理论与方法》，有改动

敏感性问题的抽样调查

当调查的内容涉及到政治态度或个人隐私等敏感性问题时，被调查者往往会拒绝回答或者给出虚假的回答. 因此，对这一类问题，必须采用经过特别设计的调查方法，以消除被调查者的顾虑，使他们做出真实的回答. 下面介绍两个方法：

（1）沃纳随机化回答模型

这种模型是由沃纳（Warner）首先提出的. 它是向被调查者显示两个与敏感性问题有关但完全对立的问题. 例如，某中学为了正确估计本校学生在考试中曾有过作弊行为的人占有的比例，随机抽取 n 个学生进行调查，向每个学生出示两个问题：

问题 1　你在考试中作弊过，是吗？请在下面表格相应处打√：

（　）是	（　）否

问题 2　你在考试中没有作弊过，是吗？请在下面表格相应处打√：

（　）是	（　）否

事先准备好一叠完全相同的卡片，上面分别写着 1 或者 2，写 1 与写 2 的卡片的比例为某一定数．让被调查者随机抽取一张卡片，如果抽到写 1 的卡片，就回答问题 1，反之，回答问题 2．然后可以通过结果，估计出作弊的人占有的比例．

要取得成功的关键在于：使被调查者确信，调查者不知道他回答的是哪个问题．

（2）西蒙斯随机回答模型

这个模型将沃纳模型中的第二个问题改为与所调查的敏感性问题完全无关的另一个非敏感性问题．例如

问题 1　你在考试中作弊过，是吗？请在下面表格相应处打√：

（　）是	（　）否

问题 2　你是一月份生的吗？请在下面表格相应处打√：

（　）是	（　）否

在这个随机回答模型中，我们要求具有第二个问题所指的特征人占的比例已知，比如这里的比例为 1/12（一年中有 12 个月）．

在这两个随机回答模型中，我们得到的答案都只有两种："是"与"不是"．答"是"的人有一部分真正在考试中作过弊，而另一部分人却未必在考试中作过弊（比如在沃纳随机回答模型中是没有作过弊的，在西蒙斯随机回答模型中是与此问题无关的）．因此我们有必要用统计的方法把真正作过弊的人所占的比例估计出来．这个统计方法在这里就不讲了．

12.3 用样本分布估计总体分布

12.3.1 频率分布表

教材线索

教材首先给出频率的定义，然后通过一个实例介绍频率分布表，最后给出制作频率分布表的基本步骤.

教学目标

1. 理解频率的定义.
2. 能分析和绘制频率分布表.
3. 能根据经验公式和具体实际合理地选择分组.

教材分析

1. 重点

根据样本数据绘制频率分布表.

2. 难点

(1) 如何根据样本数据合理地选择分组，确定组距以及各区间的上、下限. 使得绘制的频率分布表简单明了.

(2) 对频率分布表进行简单的分析.

3. 实际中，我们得到的样本数据量是比较大的，我们所需要的一些有用的信息往往淹没于纷繁的数据之中. 这些数据通常是杂乱无章的，只有通过整理才能从中发现规律. 数据整理的一种常用的方法是给出数据的频率分布表.

4. 频率（frequency）的理解

教材中关于频率的定义是描述性的.

记 f_i 为观测数据 y_i 的频率，则 $f_i=\frac{n_i}{n}$，其中 n_i 为数据 y_i 出现的次数，称之为频数. 从定义中可以得出频率的如下简单的性质：

(1) $0<f_i\leqslant 1$；

(2) $\sum f_i=1$，即所有数据出现的频率之和为 1.

5. 下面给出制作频率分布表的基本步骤

第一步：将数据从小到大排列，找出数据中最大值和最小值，计算两者之差，称之为极差，

记作 R. 再根据样本量 n 对数据进行分组，通常可以分为 5～15 组，组数 k 有一个经验公式：

$$k=1+4\lg n.$$

上述 k 的公式不是绝对的，可以有较大的灵活性，只要能显示数据的规律即可.

注：有的统计书上提供的经验公式为：$k=1+3.3\lg n$.

第二步：确定组距 d. 一般采用等间隔分组，可以根据分组数 k 和极差 R 确定组距 d：

$$d=R/k.$$

其中第一组的左端点要比数据的最小值小，最后一组的右端点要比数据的最大值大.

第三步：确定各区间的上、下限. 要使得每个数据只能落到一个组区间内，这时可以采用组区间的上、下限值比原始数据的测量精度高一位，或者用半开半闭等办法. 绘制频率表的第一列.

第四步：计算出各组的频数 n_i，填入表的第二列.

第五步：计算各组的频率 f_i，填入表的第三列.

第六步：汇总，将第二、三两列的和填入最后一行 （这一步可省略）.

例. 表 12.1 是上海市中心气象台 99 年（1884－1982 年）的上海年降水量的资料（单位：mm）（本例出自《统计手册》）.

表 12－1 上海市年降水量（单位：mm）

（1884－1982 年）

1184.4	1113.4	1203.9	1170.7	975.4	1462.3	947.8	1416.0	709.2
1147.5	935.0	1016.3	1031.6	1105.7	849.9	1233.4	1008.6	1063.8
1004.9	1086.2	1022.5	1330.9	1430.4	1236.5	1088.1	1288.7	1115.8
1217.5	1320.7	1087.1	1203.4	1480.0	1269.9	1040.2	1318.4	1192.0
1016.0	1508.2	1159.6	1021.3	986.1	794.7	1318.3	1171.2	1161.7
791.2	1143.8	1602.0	951.4	1003.2	840.4	1061.4	958.0	1025.2
1265.0	1196.5	1120.7	1659.3	942.7	1123.3	910.2	1398.5	1208.6
1305.5	1242.3	1572.3	1416.9	1256.1	1285.9	984.8	1390.3	1062.2
1287.3	1477.0	1017.9	1217.7	1197.1	1143.0	1018.8	1243.7	909.3
1030.3	1124.4	811.4	820.9	1184.1	1107.5	991.4	901.7	1176.5
1113.5	1272.9	1200.3	1508.7	772.3	813.0	1392.3	1006.2	1108.8

分析：本例数据较多，在分组、组距上采用较为灵活的方法. 最小值为 709.2，最大值为 1 659.3，极差为 950.1. 取 $k=11$，$d=100$，具体分组及各组的频数，频率列于表 12.2.

表 12－2 年降水量频数频率分布表

组号	区间	频数	频率
1	(620, 720]	1	0.010 1
2	(720, 820]	5	0.050 5
3	(820, 920]	6	0.060 6
4	(920, 1 020]	17	0.171 7
5	(1 020, 1 120]	18	0.181 8
6	(1 120, 1 220]	22	0.222 2
7	(1 220, 1320]	14	0.141 4
8	(1 320, 1 420]	7	0.070 7
9	(1 420, 1 520]	6	0.060 6
10	(1 520, 1 620]	2	0.020 2
11	(1 620, 1 720]	1	0.010 1

从上表可以看出，在 99 年中有 78 年的年降雨量在 920 mm 和 1 420 mm 之间，占总数的 78.76%，超过 1 520 mm 的仅有 3 年，占总数的 0.030 3%，低于 820 mm 的也只有 6 年，占总数的 0.060 6%.

若数据没有按从小到大排列，在确定频数时可以采用唱票的方法统计.

教学建议

(1) 在统计学中，图形和表格占有相当重要的地位. 变量间的联系，大自然中的各种因果关系，某一事物的发展变化趋势等等，光从数据上可能很难发掘出一些有用的东西，但一旦制成各种图表，可能这种联系或者趋势就能直观地显示出来了. 能“看图说话”或者“看表说话”进而自己能绘制各种图形表格，是统计工作者的一项基本技能. 从本节开始，学生就开始接触各种图形和表格，教师在讲授时要强调图表在统计中的作用.

(2) 合理的分组，确定组距和各组区间的上、下限乃是绘制好频率分布直方图的关键，也是后面学习频率分布直方图以及频率分布折线图的基础. 教师应多结合实际讲解几道例题，使学生能掌握该方法.

相关链接

统计学中的盐

1947 年印度刚独立，德里就发生了一些公共暴乱. 一个少数民族团体中的大多数人避难到被称为红色堡垒的地方，这是一个被保护的区域. 少部分人逃到另外一个地方的修姆因庙里，这个庙临近的一个古建筑物. 政府有责任提供食物给这些避难者. 这个任务委托给了承包商，由于没有任何关于避难人数的信息，政府被迫接受和付出承包商所提出的为避难者所购买的各种日用品和生活保障品的帐单. 政府的这项开支看起来非常大，因而有人建议让统计学家（他们能计算）来求出红色堡垒中避难者的正确人数.

在当时的混乱条件下，这个问题看起来很困难. 另一个复杂的情形是，政府所请的统计学家是属于多数派团体的（与避难者所属团体对立），因而如果要应用统计技术估计避难者的人数而要求进入红色堡垒的话，这些统计专家的安全没有保证. 摆在统计学家面前的问题是：在没有任何避难者人数的先验信息、没有任何机会直接了解那个地区人口密度的情形下，同时在不能使用任何已知的用于估计或人口统计调查的抽样技术条件下，来估计一个给定地区的人口数量.

专家们不得不想出某个办法来解决这个问题. 无论是统计学或是统计学家的失败，政府都是容忍的. 不管怎样，统计学家们接受了承包商交给政府的帐单，这些帐单记载了提供给避难者的不同的生活用品，如所购入的米、豆类和盐. 如何利用这些资料呢？

假设全体避难者一天所需要的米、豆类和盐的总量为 R，P，S. 由消费调查，每人每天所需要的这些食物的量分别设为 r，p，s. 因而 R/r，P/p，S/s，提供了一个集团中相同人数的平均估计量. 也就是说，这三个值无论哪一个都是等价有效的. 专家们利用承包商提出的 R，P，S 计算了这些值，发现 S/s 最小，而表示大米的 R/r 最大. 与盐相比，商品中最贵的大米的量可能被夸大了（当时在印度盐的价格非常低，因而不会夸大盐的用量）. 因此，统计学家提出估计值 S/s 为红色堡垒中避难者的人数（这里的人数要少得多），得到了很好的近似值.

由统计学家所给出的估计值对政府做出行政管理决策时非常有用. 这也提高了统计学的威信. 从那以后，统计学受到政府的大力支持. 可以说，这个估计方法对印度统计学的发展做出了很大的贡献.

这里所用的方法在任何教科书中都没有记载，是一个非惯例而且是很巧妙的方法. 这个思想的背后是统计的推理或定量的思考，或许也可以说包含了一种艺术成分吧.

——摘自《统计与真理——怎样运用偶然性》

12. 3. 2 频率分布直方图

教材线索

本小节紧紧与上小节衔接，将上小节的频率分布表进一步绘制出频率分布直方图，使得频率分布更加形象和直观.

教学目标

会根据数据绘制频率分布直方图.

教材分析

1. 重点

频率分布直方图的绘制.

2. 难点

如何从频率分布直方图中获取有用的信息.

3. 直方图能直观地显示数据的规律性

在平面坐标上，以 x 轴表示所考察的数据变量，并标出各区间上、下限，以 y 轴表示频率/组距. 以每一组的区间为底，以该区间上的频率/组距为高画一个长方形，便可画出数据的频率分布直方图. 只要改变纵轴单位，可使得频率分布直方图的形状不变.

下面例子给出上节上海市年降水量的频数与频率分布直方图.

例. （续前例）上海市年降雨量频率分布直方图.

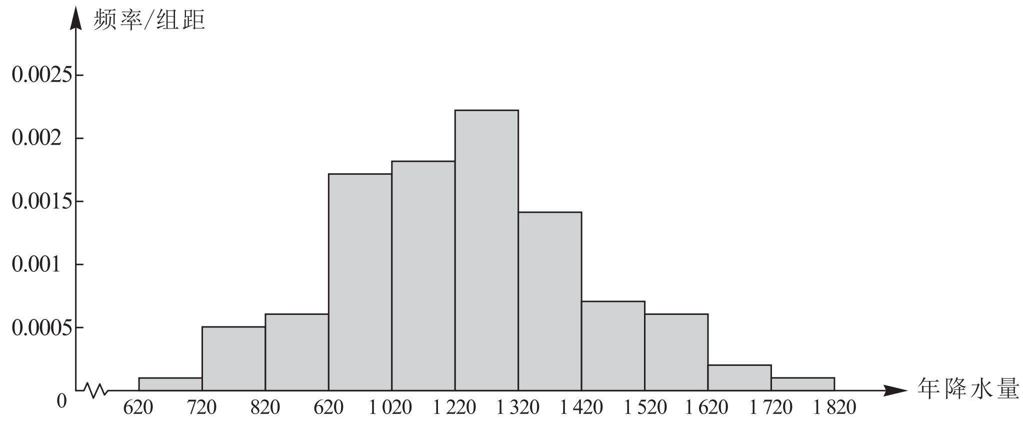


图 12－1 上海市年降水量频率分布直方图

5. 频率分布直方图

频率分布直方图能大致反映样本的分布情况，样本是从总体中随机抽取的，因而频率直方图可以推测总体的大致分布情况.

教学建议

本小节是在上小节频率分布表的基础上学习的，教师可以适当复习怎样绘制频率分布表. 频率分布表的分组数，组距与组区间确定了，绘制直方图也就简单了.

例题分析

例 1. 根据下表的数据，绘制频率分布直方图.

重量（克）	频数	频率
13	2	0.022 2
14	2	0.022 2
15	4	0.044 4
16	18	0.200 0
17	24	0.266 7
18	35	0.388 9
19	5	0.055 6
共计	90	1.000 0

解： 本例我们不需要对原始数据进行分组，所以可以直接绘制频率分布直方图如图 12－2 所示：

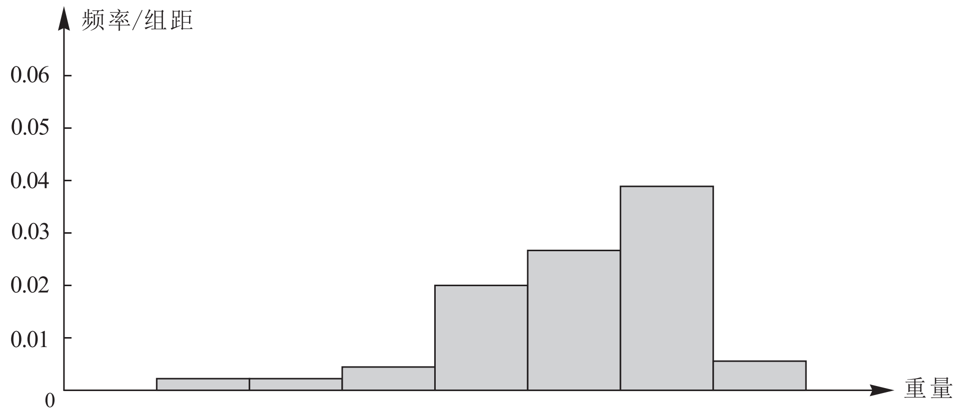


图 12－2

例 2. 某校为了了解学生的课外阅读情况，随机调查了 50 名学生，得到他们在某一天各自课外阅读所用时间的数据，结果用 12－3 所示的条形图表示. 根据条形图可得这 50 名学生这一天平均每人的课外阅读时间为 （　　）

A. 0.6 小时

B. 0.9 小时

C. 1.0 小时

D. 1.5 小时

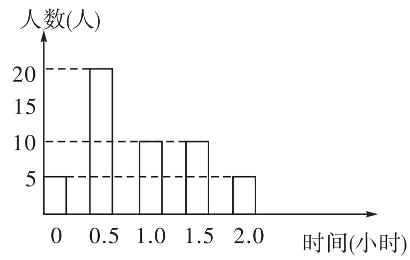


图 12－3

解： 从直方图可以看出，在这一天中，50 名学生中有 5 人不阅读，20 人阅读 0.5 小时，有 10 人 1.0 小时，有 10 人阅读 1.5 小时，有 5 人 2 小时. 因此，这 50 名学生在这一天平均每人的课外阅读时间为

$$\frac{5\times0+20\times0.5+10\times1+10\times1.5+5\times2}{50}=0.9\text{（小时）}$$

所以答案应该为 B.

例 3. 一个社会调查机构就某地居民的月收入调查了 10 000 人，并根据所得数据画了样本的频率分布直方图（如图 12－4）. 为了分析居民的收入与年龄、学历、职业等方面的关系，要从这 10 000 人中再用分层抽样方法抽出 100 人作进一步调查，则在 [2 500，3 000)（元）月收入段应抽出____________人.

解： 从频率直方图来看，月收入在 [2 500，3 000)（元）区间内的居民的频率为 $0.000\,5\times500=0.25$，现在要再抽出 100 人作进一步调查，则应该在此收入段抽出 $0.25\times100=25$ 人进行调查.

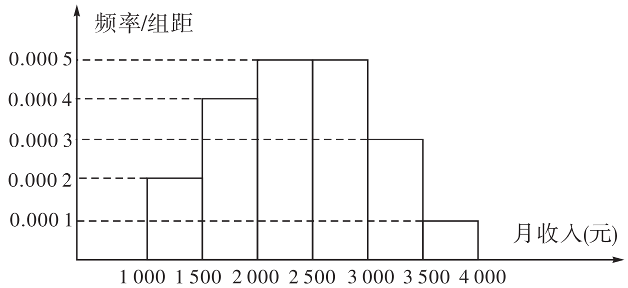


图 12－4

相关链接

一、直方图一词的起源

单词“histogram”（直方图）第一次被使用是在 1895 年，由伟大的英国统计学家卡尔・皮尔逊在脚注里作了定义. 在他给伦敦的皇家协会发表的讲话中，皮尔逊提到一些关于在 1885—1886 年英格兰和威尔士地区的房地产估价的数据：“到观测能被画成理论曲线为止，看来是不能指望有什么结果的. 然而直方图却显示了曲线末端的偏差数量.”皮尔逊没有解释他为什么使用这个特殊的术语.

比较语言学家 Eric Iversen 给出了单词“histogram”一个学术上的解释：

“histogram”从词的后面开始，可有一个比较简单的解释. “gram”当然是指图或代表什么东西，如 pictogram（彩色照片）——涂色的画像；telegram（电报）——来自远方的话语；epigram（警句）——和某种东西有关话语. 你可以看到“gram”可以代表话语或图象，其意义依赖于你怎样使用它. 但是“hist”更有趣. 由于它有多种含义，因此可能会有人误解它，象“history（历史）”中，希腊语词根“historia”；而“histology（组织学）”——关于人体组织的研究中，就是拉丁语“hist”，表示连接或组织. 但是“histogram”中的“histo”是希腊语，表示桅杆或大梁. 我想，之所以用“histogram”，仅仅是因为图中的列看起来象船上的桅杆或大梁.

——摘自吴喜之等译：《统计学　基本概念和方法》

二、具有欺骗性的直方图

在设计频率频数分布直方图时，一定要遵循一些科学的原则. 有些统计工作者为了达到某种目的，设计出来的直方图可能会产生某种假象，从而达到欺骗的目的. 请看下例：

下届市长选举中有三个候选人：A、B 和 C. 为了了解候选人受欢迎的程度，抽取了 100 名选举人进行民意测验. 图 12－5 为候选人 C 构造的民意测验结果的频数直方图，请问每个候选人的支持者分别为多少？为什么这个直方图不是测验结果的真实表示，具有欺骗性？

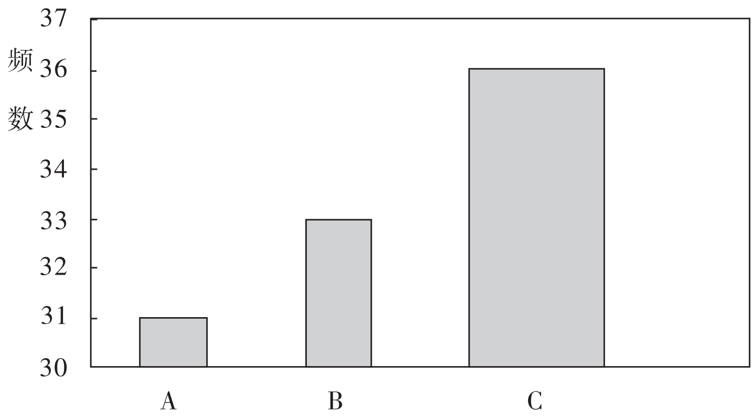


图 12－5

分析：民意测验的结果非常接近，候选人 A 的得票为 31 票，候选人 B 的得票为 33，候选人 C 的得票为 36．但是从直方图分析，似乎候选人 C 获得了胜利．这里有三个迷惑的地方：不从 Y 轴的零点开始，而是从 30 开始，从而夸大了矩形间的差别；横轴为纵轴的 75％，使得矩形显得越发细长；用来表示候选人 C 的矩形的底宽是其他矩形的两倍，使得矩形的面积非常大．所以说这个直方图具有欺骗性．

12.3.3　频率折线图

教材线索

教材首先给出频率折线图的画法，然后结合一个有趣的例子来分析频率折线图．

教学目标

了解频率折线图，会分析和绘制频率折线图．

教材分析

1．重点

绘制频率折线图．

2．难点

分析频率折线图．

3．本小节是上小节频率分布直方图的延续． 学生学会了分析和绘制频率分布表与频率分布直方图以后，对本小节的内容不难掌握．

4. 与频率直方图相比，频率折线图与密度函数曲线更接近. 若我们用一条光滑曲线连接每个矩形上边的中点，则当样本量较大，组距较小时，我们一般将这条曲线当作总体分布密度函数的一个估计. 频率折线图实际上是总体密度函数的一个自然估计（Naive estimate）. 若我们让每一组的组距不再固定，而是随着数据变化，则得到的曲线就是密度函数的核估计（Kernel estimate），这是当今统计学中一个非常流行的非参数方法. 限于学生的理解水平，教师可以不必做过多介绍.

例.（续前例）上海市年降雨量的频率折线图，如图 12－6.

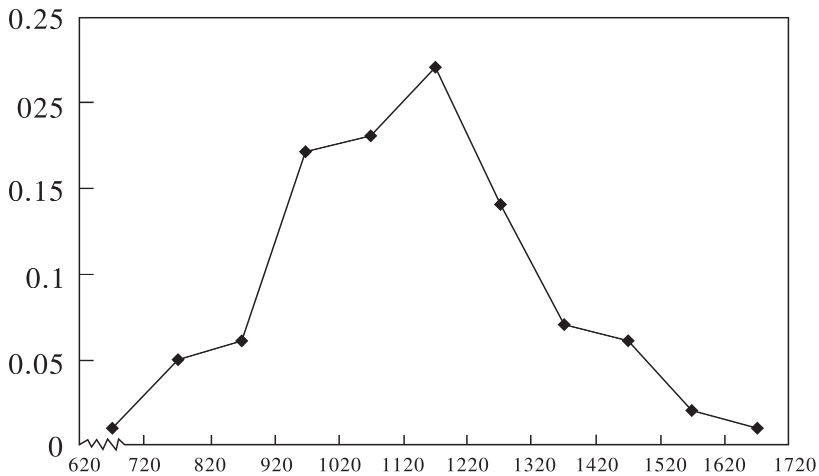


图 12－6　上海市年降水量频率折线图

教学建议

本小节的内容是与前两节内容紧密相连的，教师教授时可对上两节内容进行简单复习.

相关链接

统计资料的图形化

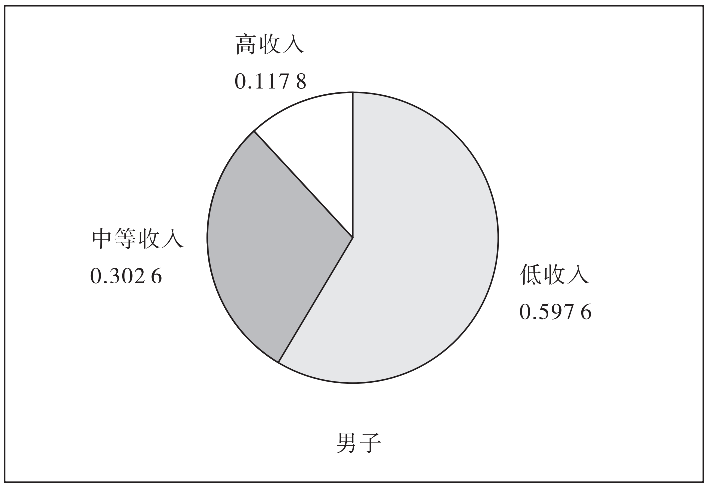
为了可以对采集到的数据进行统计分析，通过各种图形把数据的分布情况直观地反映出来是很有益的. 教材中已经介绍了直方图，折线图和茎叶图，下面再介绍一种常见的图——饼状图（或称圆形图）.

饼状图是把全圆分成若干扇形部分，以各扇形面积的大小表示频率或百分数的值. 下面以一例来说明之.

例. 下面数据来自《美国统计摘要》（1995），表中分别给出了低收入、中等收入以及高收入的男子单身家庭和女子单身家庭的频数和频率，试分别作出男子和女子的频率饼状图.

家庭收入	单身家庭数（以 1 000 为单位）			
	男子单身		女子单身	
	频数	频率	频数	频率
低收入	5 471	0.597 6	10 795	0.761 7
中等收入	2 856	0.302 6	2 661	0.187 8
高收入	1 112	0.117 8	716	0.050 5
共　计	9 439	1.000 0	14 172	1.000 0

解：频率饼状图如下：



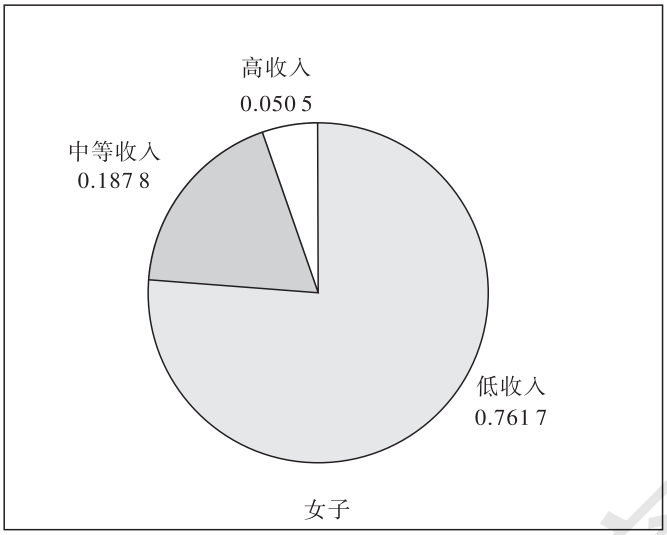


图 12－7

——节选自伯恩斯坦等著《统计学原理——描述性统计学与概率》

12.3.4 数据茎叶图

教材线索

教材首先介绍茎叶图的特点与作用，然后通过几个例子介绍怎样制作茎叶图，最后总结了茎叶图的优缺点.

教学目标

1. 会制作数据的茎叶图.
2. 能通过茎叶图了解数据的分布情况.

教材分析

1. 重点

（1）茎叶图与双茎叶图的制作.

（2）通过数据的茎叶图分析数据的分布特征.

2. 难点

（1）茎叶图的“茎”和“叶”的确定. 双茎叶图的制作.

（2）通过茎叶图分析数据的分布情况.

3. 数据的茎叶图

为了展示一批数据的特征，把数据用茎和叶所组成的数字图形来表达数据的分布情况，这种图形称为茎叶图. 教材已经通过两个例子介绍了如何制作和分析茎叶图与双茎叶图，下面再通过一个实例来说明茎叶图的制作方法.

例. 1999 年 7 月 1 日至 7 月 23 日上海市每天的日最高温度（单位：摄氏度）依次为：

26.1	24.8	29.1	28.9	26.6	25.1	25.4	30.6
28.1	25.9	26.6	24.5	25.9	28.9	26.4	27.5
25.1	25.3	27.6	30.6	30.3	32.0	33.6	

把每个数据分成茎和叶两部分，如 26.1 可把 26 作为茎，1 作为叶. 将所有茎从小到大列于茎叶图的左侧，右侧叶尾随其后，就得到 23 个数据的茎叶图，如图 12－8 所示. 从茎叶图中可知茎 24 有 2 张叶子（即 8 和 5），它们表示原始数据中有两个数据，即 24.8 和 24.5，其余类推. 图中有茎而没有叶，说明原始数据中没有这个数，例如 31 这根茎就是这个情况. 把每根茎上的叶从小到大地排列，得到如图 12－9 的茎叶图，这两张图的效果是一样的.

茎	叶
24	85
25	149913
26	1664
27	56
28	919
29	1
30	663
31	
32	0
33	6

图 12－8

茎	叶
24	58
25	113499
26	1466
27	56
28	199
29	1
30	366
31	
32	0
33	6

图 12－9

从上图可以看出大约有三分之二的天中最高温度在 25.1 ℃至 29.1 ℃之间. 有 2 天超出 31 ℃的天气，反映了数据的分布情况.

茎叶图中数据的茎和叶的划分，要根据数据的特点灵活地决定. 例如，若给出的一批数据是由二位数组成，此时可以把十位数作为茎，个位数作为叶. 若给出的数据有一位小数，则可将个位以上的数作为茎，小数点以后的数作为叶，其余情况类推.

4. 数据的双茎叶图

在同一个茎叶图上表现两组数据的分布情况，这样做可以有利于两组数据进行比较.

5. 茎叶图的优缺点

优点：（1）可以很直观地看出数据的分布情况. （2）双茎叶图还有利于对两组数据进行比较.

缺点：（1）数据量很大时，茎叶图效果不好. （2）特别分散的数据不宜制作茎叶图.

教学建议

茎叶图可以直观地反映出数据的大致分布情况，双茎叶图还可以对两组数据进行比较，因此茎叶图也得到了比较广泛的应用. 教师在教学中既要讲授制作茎叶图的方法，更重要的是要使学生学会分析茎叶图. 能透过各种统计图表汲取对我们有用的知识，是一种能力，也是每一个统计工作者所必备的素质. 教师应加强对学生的这种能力和素质的培养.

例题分析

例 1. 某班 40 名学生的数学考试成绩按从小到大排列为：48，55，60，61，62，62，67，69，70，71，71，72，73，74，74，75，75，75，75，76，76，77，78，79，80，80，81，83，83，84，85，85，86，87，87，88，90，93，95，98.

请制作数据的茎叶图并进行简单的分析.

解：以十位数为茎，个位数为叶制作茎叶图如下：

茎	叶
4	8
5	5
6	012279
7	0112344555566789
8	001334556778
9	0358

图 12－10

从图中可以直观地看出，该班数学成绩分布范围为 48 分至 98 分，其中七十多分和八十多分的占绝大多数，成绩优秀的（90 分以上）的比较少，不及格的也比较少.

相关链接

五数概括

我们除了可以用频率分布直方图、频率折线图以及数据茎叶图等图形来直观地反映数据的分布情况．下面介绍五数概括，它也能反映数据的分布情况.

设有容量为 n 的一组数据（样本）$\{x_1, x_2, \cdots, x_n\}$，将其按从小到大排列后记为 $\{x_{(1)}, x_{(2)}, \cdots, x_{(n)}\}$，称排列好以后的数据为原始数据的有序数组．下面给出该数据的数字特性.

1. 中位数　一组数据（样本）按从小到大排列以后，中间位置的数称为数据的中位数．用字母 M 或者符号 $\tilde{x}$ 表示．依据 n 的奇偶性，中位数的计算公式为：

$$M=\begin{cases} x_{(k)}, & n=2k-1, \\ \dfrac{x_{(k)}+x_{(k+1)}}{2}, & n=2k. \end{cases}$$

可见，当 n 为奇数时，中位数恰好是有序数组中的中间一个数；当 n 为偶数时，中位数不是有序数组中的一个数，而是中间两个数值的平均.

2. 极值　它有两个，最小值 $x_{(1)}$ 和最大值 $x_{(n)}$.

3. 四分位数　它也有两个：一个在中位数和最小中的中间，称为下四分位数，记作 F_L；另一个在中位数和最大值的中间，称为上四分位数，记作 F_U.

把中位数，上、下四分位数和最小、最大极端值这 5 个数放在一起，已能刻画出这组数据的大致状态，称为五数概括．由五数概括可看到：

这组数据的中心在何处？

这组数据的分布在什么范围内？

这组数据的中段分布在什么范围内？

例如，中位数反映了这组数据的中心位置或典型值，在中位数两侧 F_L 和 F_U 之间包含了大约一半的数据，而其宽度 F_U-F_L 表征这组数据的散布大小．在 F_L 和 F_U 的外端分别包含了各有四分之一的数据，而四分位数的名字也是由此而来．

例． 从某厂生产的一种型号的节能灯中抽出 14 个进行寿命试验，测得其寿命，按从小到大的次序排列（单位：小时）：

120　548　730　820　918　1 100　1 250　1 430　1 626　1 800　1 980　2 190　2 800　3 970

试给出这组数据的五数概括．

解： 样本量 $n=14$，中位数为

$$M=\frac{x_{(7)}+x_{(8)}}{2}=\frac{1\ 250+1\ 430}{2}=1\ 340.$$

两个四分位数分别为：

$$F_L=x_{(4)}=820,\quad F_U=x_{(11)}=1\ 980,$$

极值 $x_{(1)}=120$，　$x_{(14)}=3\ 970$．

——摘自茆诗松等：《统计手册》

12.4 数据的相关性

12.4.1 相关性

教材线索

教材首先通过一个身高与体重的关系指出成对数据可能会有相关关系存在. 然后介绍数据的散点图，最后介绍怎样通过散点图来初步判断数据间的相关性.

教学目标

理解数据间的相关性的含义，能通过数据的散点图初步判断数据间是否存在相关性，以及是何种相关（正相关还是负相关）.

教材分析

1. 重点

数据相关性的理解；通过数据的散点图分析数据间的相关性.

2. 难点

如何判断数据间存在的相关性以及相关的种类.

3. 数据间的相关性

在实际问题中，我们会经常遇到有相关关系的量. 比如教材中描述的身高和体重的关系，一般说来，身高者，体重也重，但是身高并不能完全确定体重. 这样的例子还有很多. 正相关的情况：比如气温变化与饮料的销售，一般说来，气温升高，饮料的销售量就会增加. 还有施肥与农作物产量的关系，施肥多，农作物的产量就高等等. 负相关的例子也有很多：比如说商品的价格与销售量. 价格越高，销售量越小. 价格越低，销售量越高. 但是销售量不完全由价格决定，它还与其他的因素有关，比如产品的质量，人们对品牌的信赖度等等.

从某种意义上来说，函数关系是一种理想的关系模型，它是一种确定性的关系. 即只要自变量确定了，因变量则可由函数关系完全确定. 相关关系则是一种更为一般的情况，由于有随机因素的影响，因变量不能完全由自变量确定. 研究相关关系，不仅可以使我们处理更为广泛的数学应用问题，而且还可以使我们对函数关系的认识上升到一个新的高度.

4. 数据散点图

散点图是一种非常重要的统计图，它是在同一个坐标平面里面把数据点绘制出来. 它有数

据的散点图（横坐标为一个数据，通常是自变量，纵坐标为另一个数据，通常是因变量），数据与残差的散点图（横坐标为数据，纵坐标为残差）等等. 我们通过散点图可以很直观的判断出数据间的相关关系，异方差性等. 统计工作者要学会绘制各种图形，以及从各种图表中获取有用的信息，能根据图表初步做出判断等.

5. 从数据的散点图判断相关性

高度正相关：数据点（x，y）十分明显地集中在一条上升的直线附近.

中度正相关：数据点（x，y）也分布在一条上升的直线附近，但集中的程度不十分明显.

高度负相关：数据点（x，y）十分明显地集中在一条下降的直线附近.

中度负相关：数据点（x，y）也分布在一条下降的直线附近，但集中的程度不十分明显.

注意到上面所使用的语言都是一些描述性的. "十分明显"与"不十分明显"没有明确的界定，是一种模糊语言，缺乏统计上的一种精确描述. 从这里可以看出利用散点图只能初步判断数据间的相关性，这是它的局限性所在.

至于如何将这种相关性数量化，统计学里面有两个与之联系的量：可决系数与相关系数. 其中可决系数是一个介于 0，1 之间的数，可决系数越高，数据的相关性越强. 相关系数是一个介于－1，1 之间的量，若相关系数为正数，则数据间存在正相关，反之为负相关. 相关系数的绝对值越大，数据的相关性越强. 这些系数的值为多大，我们就接受数据就存在相关性呢？这里涉及到统计的假设检验问题. 限于高中学生的知识程度与理解、认知水平，教材没有涉及到这些方面的知识. 这些知识将在大学的《数理统计》课程中遇到.

例 1. 在两个经济变量之间，例如人均收入与消费支出之间，一般说来，收入越高，消费支出就越大. 为研究这两个经济变量之间的关系，我们统计了某市从 1990 年以来的人均收入与人均消费支出的七组数据：如下表所示

表 12.3　　　　　　　　　　　　　　　　单位：10 元

年份	人均收入 x	人均消费 y	年份	人均收入 x	人均消费 y
1990	480	420	1994	640	580
1991	510	450	1995	700	620
1992	545	490	1996	760	680
1993	590	530			

在上表中，以 x 表示人均收入，y 表示人均消费支出. 从表中数据可见，x 和 y 之间的变化，呈现出某种规律性，即随着人均收入的提高，消费支出也相应地增大. 把这些数据点绘制成散点图，如图 12－11，就可以直观的看到，这些数据点散布在一条上升的直线两旁.

6. 关于绘制数据的散点图

学生可以通过描点作图的方法绘制出数据的散点图. 但是这种方法有较大的局限性. 当数据比较多且复杂时，描点作图的方法就很难凑效了. 这时，我们一般利用统计软件的作图功能在计算机里面实现. 比如在著名的统计软件"S－PLUS"中，可以通过一个简单的命令

"plot（x，y）"绘制出数据对（x，y）的精确的散点图.

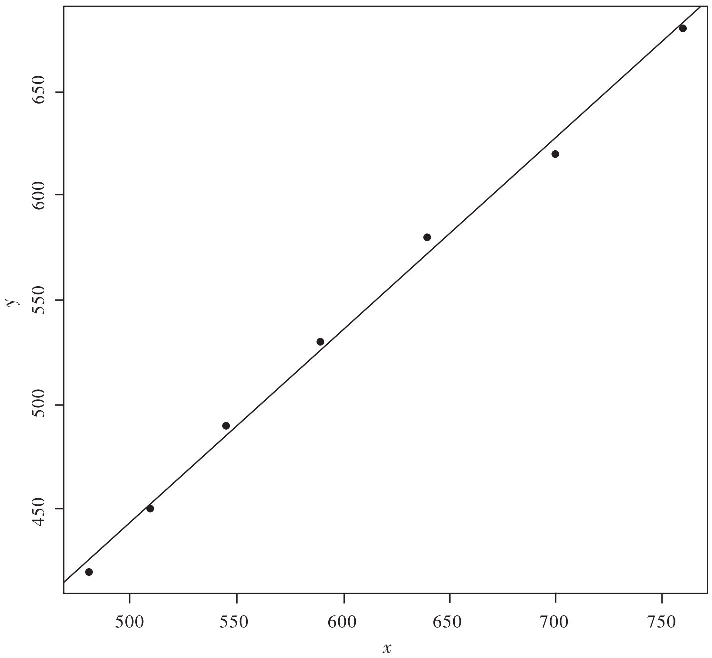


图 12－11

教学建议

1. 建议教师从丰富的实例中引入相关性的概念.

2. 统计软件"R"可以说是"S－PLUS"的免费版，非常受到国内外统计学家的青睐与推崇．目前它的最高版本为 2.3.1. 它可以从网络上免费下载.

有条件的学校，教师可以下载安装此软件，就可以利用它的作图功能绘制出精确的散点图. 让学生直观的看到数据间的相关关系. 同时也可以激发学生的学习兴趣. 例如要绘制出上面例 1 的散点图，可以在命令窗口中输入如下简单代码，就可以得到上图.

```
x<－c（480，510，545，590，640，700，760）
y<－c（420，450，490，530，580，620，680）
plot（x，y）
```

若要得到图中的那条直线，只需增加以下命令：

```
z<－lm（y～x）
abline（z）
```

若要得到更多的关于该统计软件的知识，可以阅读相关教程或者帮助文档.

例题分析

例 1. 某市某种商品的需求量同当地农村的人均月收入有关，历史上的需求与收入的数据如

下表所示.

样　本　号	1	2	3	4	5	6	7	8
收入（元）	350	400	450	500	550	600	630	670
需求（万元）	45	48	51	58	62	65	69	75

请将收入当作自变量，需求作为因变量，绘制数据的散点图，并分析数据间的相关关系.

解： 将收入作为横坐标，需求作为纵坐标，绘出数据的散点图如图 12－12 所示. 从散点图中我们可以看出，数据点明显地集中在一条上升的直线附近，因此我们可以判断出数据间存在比较强的正相关关系.

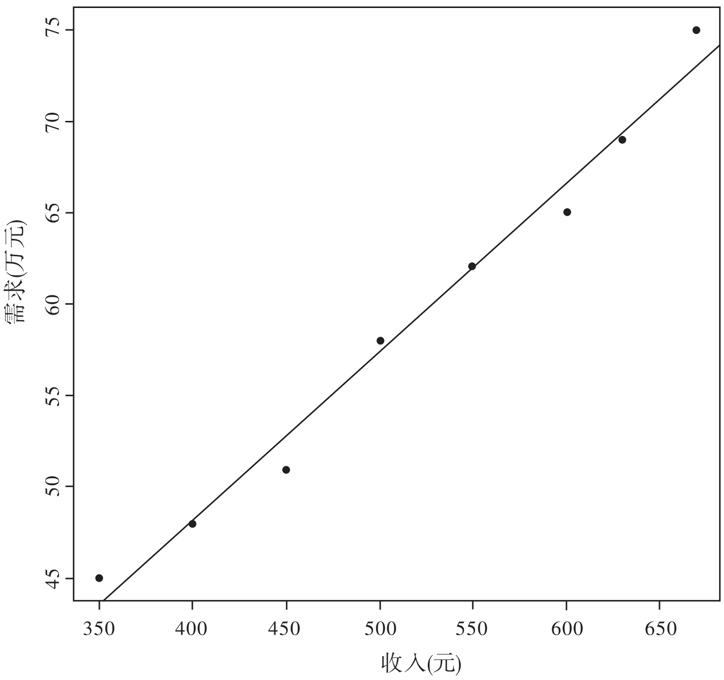


图 12－12

例 2. 某种商品的需求量与价格有关，由于原材料紧张，不能保证供给，只好用价格来控制. 价格与需求的关系有如下数据.

价　格（x）	1.5	1.8	2.0	2.1	2.3	2.5	2.8
需求量（y）	35	32	30	29	27	24	21

请将价格（x）作为自变量，需求量（y）作为因变量，绘制数据的散点图，并分析数据间的相关性.

解： 绘出数据的散点图如 12－13 图所示 . 从散点图中我们可以看出，数据点明显地集中在一条下降的直线附近，因此我们可以判断出数据间存在比较强的负相关关系.

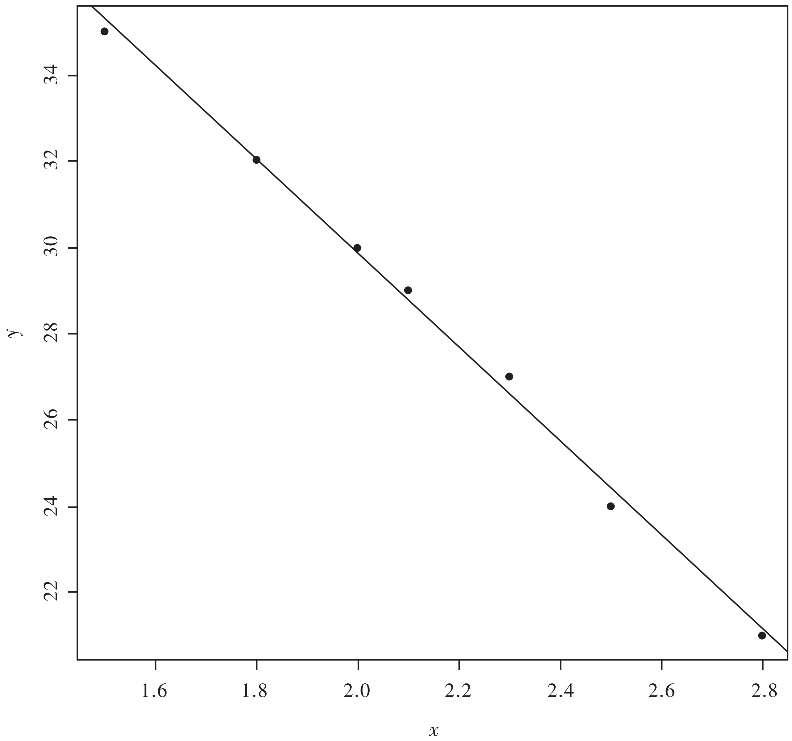


图 12-13

相关链接

趋向中间高度的回归

回归这个术语是由英国著名统计学家 Francis Galton 在 19 世纪末期研究孩子们及他们的父母身高时提出来的．正如我们所预料的，Galton 发现身材较高的父母，他们的孩子也比较高．但是这些孩子平均起来并不像他们的父母那样高．对于较矮父母情形也类似：正如我们所预料的那样，他们的孩子比较矮，但这些孩子的平均身高要比他们的父母的平均身高高（这是很自然的情况，因为如果身材高的父母他们的孩子比他们更高，而身材矮的父母，他们的孩子比他们更矮，那么许多代以后，人们的高矮差距会越来越大）．Galton 把这种孩子们的身高向中间值靠近的趋势称之为回归效应，而他发展的研究两个数值变量的方法称为回归分析．

作为老师，我们在我们的班上也看到了同样的回归现象．期中考试成绩比较高的好学生在期末考试成绩也好，但平均不如期中考试那么好．类似地，期中考试成绩比较差的学生期末考试时平均成绩要好一些．在体育上，回归效应也是一种很突出的现象：第一年成绩很突出的新手，第二年成绩往往就不是那么好．

——摘自吴喜之等译：《统计学基本概念和方法》

12.4.2 回归直线

教材线索

教材首先引入回归直线的含义，然后引出找到这条回归直线必须遵循的原则，最后利用使各点到直线距离平方和达到最小的准则求出回归直线，再通过一些例子例解如何具体求出回归直线，在本小节的最后引出参数与最小二乘估计的概念.

教学目标

1. 会根据数据求出回归直线，并进行简单的预测.
2. 体会最小二乘估计所蕴含的统计思想.
3. 了解参数的含义.

教材分析

1. 重点

利用最小二乘估计求出回归直线，并进行简单的预测.

2. 难点

（1）最小二乘估计的理解以及其所蕴含的丰富的统计思想.

（2）根据数据求出回归直线.

（3）统计预测的合理性.

3. 本小节主要研究的是回归分析中最简单、也是最基本的一种类型：一元线性回归分析. 这类似于代数中的一元一次方程. 所谓“一元”，是指因变量只与一个自变量相关. 若因变量与多个自变量相关，则相应的统计分析称为多元回归分析. 所谓“线性”，是指因变量和自变量的关系是线性关系，线性关系是最简单的一种关系. 若因变量与自变量的关系是一种非线性关系，比如平方，三角函数等，则称相应的统计分析为非线性回归分析. 最广泛的一种情况，因变量与自变量的关系是未知的，此时我们称相应的统计分析为非参数回归分析. 目前，线性回归与非线性回归已经被研究得很透彻了，而非参数回归分析则是现代统计学中的一个热点问题. 从这里我们可以看出，一元线性回归是最简单的一种回归分析，这也是符合高中学生的认知特点的.

4. 本小节主要介绍如何寻求回归直线，所借助的工具是最小二乘估计. 最小二乘法是一个古老的方法，它从诞生到现在已有几百年的历史. 最早可以追溯到 19 世纪初. 在整个统计学中

的影响是相当深远的. 可以说最小二乘法的思想一直贯穿统计学的发展，直至目前，最小二乘法一直焕发出它的青春活力.

5. 从上小节知道，若数据的散点图集中在一条直线的附近，此时数据是具有相关性的，这条直线就是回归直线. 但是这条直线是未知的，我们必须把它找出来. 要找出这条直线，必须有一个准则. 若观测数据只有两对，两点确定一条直线，显然连接这两点的一条直线就是所要找的回归直线. 但是这在实际问题中是毫无意义的. 教材通过三对观测数据，找到一个准则，即这条直线必须满足使得各数据点到这条直线的距离的平方和最小. 这就是最小二乘的思想，这里所谓的“二乘”就是平方的意思，“最小”就是使得这些平方和达到最小. 这是一个很自然的想法.

一般地，假定有样本量为 n 的一组数据：

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots,\ (x_n,\ y_n),$$

通过散点图，发现这些数据有明显的相关关系，要求出回归直线 l：$y=a+bx$，其中 a 称之为截距（intercept），b 称之为斜率（slope），这里 a 和 b 均是未知的，称之为参数. 目的是要求出 a 和 b 的值，这样回归直线就确定了. 采用下面的准则，使得

$$Q(a,b)=\sum_{i=1}^{n}(y_i-a-bx_i)^2=\min!$$

上式中 $=\min$ 的意思就是使该式达到最小的意思，引入这些记号和求和符号 $\sum$ 可以使表达式显得简单. 从上式中求得的 a 和 b 的值，称为 a 和 b 的估计，这个方法就是著名的最小二乘法.

6. 有没有其它的准则来寻求回归直线呢？

答案是有的！教材指出，采用使得绝对值之和最小的准则也是可行的，即使得

$$Q(a,b)=\sum_{i=1}^{n}|y_i-a-bx_i|=\min!$$

但是这个方法带来一些数学上的处理困难. 相应的估计方法称之为最小一乘估计. 我国已故的著名数理统计学家、中国科学院院士陈希儒老先生和他的学生们在最小一乘估计方面的研究颇有建树，处于世界领先水平.

7. 如何求得参数 *a* 和 *b* 的估计呢？

限于学生的知识水平，教材中没有给出最小二乘估计的求法，只是给出计算 a 和 b 的公式. 我们可以用配方的方法求出它们的估计值，具体补充如下，供教师参考.

$$\begin{aligned}Q(a,b)&=\sum_{i=1}^{n}(y_i-a-bx_i)^2=\sum_{i=1}^{n}(y_i^2+b^2x_i^2+a^2-2bx_iy_i-2ay_i+2abx_i)\\&=\sum_{i=1}^{n}y_i^2+b^2\sum_{i=1}^{n}x_i^2+na^2-2b\sum_{i=1}^{n}x_iy_i-2a\sum_{i=1}^{n}y_i+2ab\sum_{i=1}^{n}x_i\\&=na^2-2na(\bar{y}-b\bar{x})+b^2\sum_{i=1}^{n}x_i^2-2b\sum_{i=1}^{n}x_iy_i+\sum_{i=1}^{n}y_i^2\\&=n[a-(\bar{y}-b\bar{x})]^2-n(\bar{y}-b\bar{x})^2+b^2\sum_{i=1}^{n}x_i^2-2b\sum_{i=1}^{n}x_iy_i+\sum_{i=1}^{n}y_i^2\end{aligned}$$

$$= n[a-(\bar{y}-b\bar{x})]^2 + b^2(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2) - 2b(\sum_{i=1}^{n} x_i y_i - n\bar{x}\cdot\bar{y}) + (\sum_{i=1}^{n} y_i^2 - n\bar{y}^2)$$

$$= n[a-(\bar{y}-b\bar{x})]^2 + b^2(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2) - 2b\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) + \sum_{i=1}^{n}(y_i-\bar{y})^2$$

$$= n[a-(\bar{y}-b\bar{x})]^2 + \sum_{i=1}^{n}(x_i-\bar{x})^2\left[b-\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]^2$$

$$-\frac{[\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})]^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \sum_{i=1}^{n}(y_i-\bar{y})^2.$$

在上式中，后两项与 a、b 无关，而前两项为非负数，因此要使 $Q(a,b)$ 取得最小值，当且仅当前两项的值均为零，即有

$$b=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2},\quad a=\bar{y}-b\bar{x}.$$

在推导中用到下面两个公式：

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2,$$

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\cdot\bar{y}.$$

其中第一式在初中“统计初步”里出现过，第二式也很容易证明.

若采用大学数学中利用求偏导数来求二元函数的极值的知识，则整个推导过程就要简单得多. 兹推导如下，供教师参考，学生不必了解.

分别令 $\frac{\partial Q(a,b)}{\partial a}=0$，$\frac{\partial Q(a,b)}{\partial b}=0$，解下述方程组

$$\begin{cases}\dfrac{\partial Q(a,b)}{\partial a}=-2\sum_{i=1}^{n}(y_i-a-bx_i)=-2(\sum_{i=1}^{n}y_i-na-b\sum_{i=1}^{n}x_i)=0 & (1)\\ \dfrac{\partial Q(a,b)}{\partial b}=-2\sum_{i=1}^{n}(y_i-a-bx_i)x_i=-2(\sum_{i=1}^{n}x_iy_i-a\sum_{i=1}^{n}x_i-b\sum_{i=1}^{n}x_i^2)=0 & (2)\end{cases}$$

由（1）式可得 $a=\bar{y}-b\bar{x}$，把它代入（2）式，并通过化简，可得 $b=\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$.

若记 $l_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2$，$l_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$，则我们可以得到 a 和 b 的估计的简单表达式：$b=\dfrac{l_{xy}}{l_{xx}}$，$a=\bar{y}-b\bar{x}$.

注：教材中的计算公式为 $b=\frac{s_{xy}}{s_x^2}, a=\bar{y}-b\bar{x}$，其中 $s_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ 为 $\{x_i\}$ 的方差，$s_{xy}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$.

8. 关于参数估计

在一元线性回归中，截距 a 和斜率 b 都是未知的，称为参数. 在许多统计模型里面，都有一些未知的参数. 为了能进行统计推断和预测，这些未知参数必须通过各种方法利用观测数据给大致的求出来，相应的工作称为参数估计. 参数估计是统计推断里面一项十分重要的工作，包括现代的统计学，参数估计仍然是一项主要的工作.

9. 关于预测

教材例 1 要求预测某天的冷饮销售量，预测是统计的一大重要功能. 根据过去和现在的值来推断未来的值就称为预测. 只有当数据是高度相关的，这种预测才有意义. 也就是说，这些数据内在存在着某种规律性的东西，正是这种内在的规律性，才使统计预测成为可能. 所以在进行预测之前，务必先画出数据的散点图进行初步的相关性判断，找出它们这种规律性. 如果数据间的相关性不强，也即自变量和因变量之间没有太多的因果关系，它们之间没有什么规律可言，则这种预测是不准确的，也是毫无价值的. 具体要求相关性要多大才能进行预测，这里涉及到一些较深的统计知识，在大学的《统计预测与决策》或者《经济预测与决策》课程中会学到，教师可不必作过多引申.

教学建议

本小节所授的内容在数理统计学里面占有十分重要的地位，它是回归分析中最简单最基本的内容——一元线性回归分析. 在这里初步遇到了最小二乘估计，统计预测等内容. 教师要着重讲解最小二乘法所蕴含的统计思想，以及统计预测的可能性等.

在本小节中，要求学生根据公式求出回归直线. 这里的计算量是比较大的，学生弄不好就会算错. 教师要培养学生良好的计算习惯，比如说要把中间结果都写出来，既便于检查又不容易犯错. 此外，由于计算量大，有条件的学校建议使用科学计算器或者利用计算机和统计软件进行计算.

例题分析

例 1. (续上小节例题分析的例 1)：请求出上小节例 1 中的回归直线. 若人均月收入为 700 元时，请预测需求为多少万元.

解：先计算出 $\bar{x}=518.75$，$\bar{y}=59.125$，再计算出

$$l_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=90\ 487.5, l_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=8\ 351.25$$

最后，我们可以得到

$b=\frac{l_{xy}}{l_{xx}}=\frac{8\ 351.25}{90\ 487.5}\approx 0.092$，$a=\bar{y}-b\bar{x}=59.125-0.092\times 518.75=11.4$

于是，回归直线为 $y=11.4+0.092x$.

从数据的散点图来看，人均月收入与需求量是高度正相关的. 所以可以利用这条回归直线进行预测.

当人均月收入 $x=700$ 元时，对 y 的预测值是

$$y=11.4+0.092\times 700=75.8\text{（万元）}$$

可见，当人均月收入为 700 元时，需求量大约为 75.8 万元.

相关链接

一、勒让德发明最小二乘法

勒让德（A. M. L egendre，1752—1833）是法国大数学家，在数学的许多领域，包括椭圆积分，数论和几何等方面，都有重大的贡献. 最小二乘法最先出现在他于 1805 年发表的一本题为《计算慧星轨道的新方法》的著作的附录中. 该附录占据了这本长 80 页的著作的最后 9 页，勒让德在这本书前面几十页关于慧星轨道计算的讨论中没有使用最小二乘法，可见在他刚开始写作时，这一方法尚未在他头脑中成形. 历史资料还表明，勒让德在参加量测过巴黎子午线长这项工作很久以后还未发现这个方法. 考虑到此书发表于 1805 年且该法出现在书尾的附录中，可以推测他发现这个方法应当在 1805 年或之前不久的某个时间.

勒让德在该书 72～75 页描述了最小二乘法的思想、具体作法及方法的优点. 他提到：使误差平方和达到最小，在各方程的误差之间建立了一种平衡，从而防止了某一极端误差，对决定参数的估计值取得支配地位，而这有助于揭示系统的更接近真实的状态. 的确，考察勒让德之前一些学者的做法，都是把立足点放在解出一个线性方程组上，这种做法对于误差在各方程之间的分布的影响如何，是不清楚的.

关于最小二乘法的优点，勒让德指出了以下几条：一是通常的算术平均值是其一特例，这我们在前面已指出了；第二，如果观察值全部严格符合某一线性方程，则这个方程必是最小二乘法的解；第三，如果在事后打算弃置某些观察值不用或增加了新的观察值，对正则方程的修改易于完成. 从现在的观点看，这方法只涉及解线性方程组是其最重要的优点之一（其他的重要优点包括此法在统计推断上的一些优良性质，以及其广泛的适用性）. 近年发展起来的，从最小二乘法衍生出的其他一些方法，尽管在理论上有其优点，可是由于计算上的困难而影响了其应用.

最小二乘法在十九世纪初发明后，很快得到欧洲一些国家的天文和测地学工作者的广泛使用，据不完全统计，自 1805 年至 1864 年的 60 年期间，有关该方法的研究论文约 250 篇，一些百科全书，包括 1837 年出版的不列颠百科全书第 7 版，都收进了有关这个方法的介绍. 在研究

论文中，有一些是关于最小二乘估计的计算，这涉及解线性方程组，高斯也注意了这个问题，给出了正则方程的命名并发展了解方程的消去法，即高斯消去法．但是，在电子计算机出现以前，当参数个数较大时，计算的任务还是很繁重．1858 年，英国为绘制本国地图作了一次大型的调查，参数个数为 920，样本量为 1 554，其数据处理用最小二乘法，用两组人员独立计算，花了两年半的时间才完成．1958 年我国某研究所计算一个炼铁方面的课题，涉及用最小二乘法解 13 个自变量的线性回归，三十余人用电动计算机算，夜以继日花了一个多月的时间．

勒让德的工作没有涉及最小二乘法的误差分析问题．这一点由高斯在 1809 年发表的正态误差理论加以补足．高斯这个理论对于最小二乘法之用于数理统计有极重要的意义，这一点在 20 世纪哥塞特、费歇尔等人发展了正态小样本理论后，尤其看得明显．正因为高斯这一重大贡献，以及他声称自 1799 年以来一直使用这个方法，所以在现今各种著作中，多把最小二乘法的发明权归之于高斯．即在当时，在这两位数学家之间就发生了关于优先权的争论，在数学史上知名度仅次于牛顿与莱布尼兹之间关于发明微积分的优先权的争论，直到近年还有学者著文讨论．现在一般认为，二人独立地作出了这一发明，但第一个用文字形式发表的是勒让德．

二、道德文章垂范人间

许宝騄先生

许宝騄，字闲若，出身于浙江杭州的名门世家．1910 年 9 月 1 日生于北京．2 岁时随家移居天津，8 岁去杭州．他幼时体质虚弱，但聪明颖悟，孩提时就能拼摆“益智图”．他从小受到传统的教育．5 岁至 14 岁在家从师读书，通读四书五经，涉猎四史及古文辞．11 岁时还写过以《花生姻缘》和《神花》为题的文言小说．他临摹的小楷，古朴神似，为表兄俞平伯手写的《古槐书屋词》曾刻版印行．他还善于巧妙地利用古文诗词制作灯谜．11 岁开始学习英文，两年后便能阅读英文的古典文学名著．

1924 年，许宝騄的父亲在杭州病逝．全家迁回天津，次年又移居北京．为了考中学，北京大学吴缉熙老师给他讲授代数、几何及三角，两个月后即取得显著成绩．由此引发了他对数学的兴趣，他的数学天才也开始显露出来．1925 年，许宝騄考入当时的北京名校汇文中学高中部．高中学习期间，他利用暑假继续钻研数学，还坚持学习法文．两年后他的法文达到能会话和作短文的程度．

1928 年，中学毕业的许宝騄考入燕京大学化学系学习．两年后他决定改攻数学，于 1930 年秋又考入清华大学数学系．这期间，他立下了要对人类知识总量有所增添的抱负．1933 年他以优异的成绩从清华毕业，获得理学学士学位．在清华学习期间，他不仅在数学方面学有所成，还参加了俞平伯先生组织的谷青社的活动．他会唱昆曲、会拉二胡，熟悉音律；每听一曲，不出几遍就能写出谱子，表现出良好的艺术才能．

许宝騄从清华毕业后即参加了赴英公费留学考试并被录取，但因体重不够未能成行．在香山休养的一年中，他注意营养，加强锻炼，使身体状况大为好转．此后，他到北京大学任助教，并于1936年再次通过赴英庚子赔款公费留学的考试，同年赴伦敦大学学院（University College，London）学习数理统计．1938年，许宝騄获得哲学博士（Ph. D.）学位；两年后，又得到科学博士（Sc. D.）学位．当时的伦敦大学学院是公认的数理统计研究中心．现代数理统计学的奠基者费歇（R. A. Fisher）、奈曼（Neyman）和皮尔逊（E. Pearson）等都在那里工作．那里吸引着来自世界各地的许多青年人，而"最优秀的学生是中国人许宝騄"．

1940年，许宝騄出于爱国热诚，不畏艰险绕道好望角从海路回到了抗日烽火中的祖国，到昆明西南联大任教．当时的西南联大生活十分艰苦，许宝騄身体也日渐衰弱．但他精心教学，培养青年学生．钟开莱、王寿仁等著名学者都得到过他的指导和帮助．教学之余，他继续科学研究工作，成果十分丰富；仅1941年，就有5篇学术论文发表．在此期间，他目睹政治腐败、社会混乱、人民痛苦，认识到政治的重要意义和中国革命的必然趋势，毅然参加了中共中央直接领导的政治团体——中国民主革命同盟．

1944年，许宝騄的导师、已受聘担任伯克利加州大学统计实验室主任的奈曼教授考虑到他的接班人问题．"照奈曼的看法，许宝騄的水平绝对可以与瓦尔特（A. Wald，著名数理统计学家）相媲美．他们是新一代的数理统计学家中的两个佼佼者"．1945年，许宝騄应伯克利加州大学和哥伦比亚大学的联合邀请，在这两所大学各工作一个学期．他典型的中国学者的风度，对科学研究工作的高标准要求以及解决困难数学问题的能力给美国同行以极为深刻的印象．奈曼对许宝騄在哥伦比亚大学工作后去伯克利供职一事特别关心．但是，"芝加哥大学、耶鲁大学以及哥伦比亚大学都希望得到许宝騄"．1946年秋，许宝騄和郝泰林（H. Hotelling）一起去了北卡罗来纳大学新建立的统计系．

1947年，许宝騄谢绝美国大学方面的婉留，毅然回到祖国，在北京大学任教；1948当选为中研院院士．他认真教学、勤奋研究；还积极参加文教界的民主革命运动，迎接解放．北京解放以后，他致电美国同行，表达自己的欢悦心情．他以极大的热情投入新中国的建设事业．解放初期，一些西方国家对新中国实行封锁．为了开展对苏联的学术交流，许宝騄自学俄文，先后主持校对了《数学分析简明教程》、《概率论教程》和《微分方程教程》等俄文翻译教材．不仅如此，他还帮助北京大学数学系的其他教师学习俄语，带领他们精读俄文原版的《数学分析八讲》．当时，前苏联和波兰概率论专家邓肯（E. B. Dynkin）和费兹（Fisz）曾先后访问中国，受到许先生的亲切接待．毫无疑问，许宝騄的这些工作对解放初期数学方面的教材建设、人材培养和学术交流具有十分重要的意义．

建国以来，许宝騄历任北京大学一级教授、概率统计教研室主任、中国科学院学部委员和第四届政协委员等重要职务．他身体状况十分差，终生未婚，把自己的身心全部贡献给祖国的教育和科学事业．50年代末他已身患肺结核、胃病和痔疮等多种疾病，行动很不方便，但仍坚持在住所主持讨论班和教研室工作．1963年他的肺部出现空洞，还坚持工作，同时领导着数理统计、马氏过程和平稳过程三个讨论班．"文化大革命"中，他受到不公正的对待，被剥夺了工

作的权利．尽管这时他已瘫痪，卧床不起，但仍保持旺盛的工作精神：只要上面直接有任务下达，总是力争提前完成；只要环境相对安定，就抓紧研究工作．1970 年 12 月 18 日，当他在简陋的住所溘然长逝时，人们在他的床前看到的是一叠叠的算草和一支使用多年的派克钢笔．

许宝騄的学术成就主要是在数理统计和概率论这两个紧密相关的数学领域中．

作为我国概率统计方面公认的学科带头人，许宝騄对学科的建设和发展作出了重要贡献．在 1956 年周恩来总理主持制定的全国科技发展规划中，概率统计和计算数学、微分方程一起被列为数学科学的重点发展方向．许宝騄参与组织领导了概率统计方面规划的制定和落实．在他的主持下，从北京大学及中山大学、南开大学等兄弟院校的数学系抽调了 50 余人，作为国内第一届概率统计专业的学生集中到北大学习．还从中国科学院数学研究所、中山大学等单位调集了一批教师，和北大的教师一起，开设出测度论、极限定理、数理统计和马氏过程等一系列专门化课程．这一有力的措施，为我国概率统计学科培养了一批教学科研骨干．同时，在他的建议和倡导下，与前苏联、东欧的学术交流一度也比较活跃，对学科的发展起了有益的推动作用．

许宝騄始终关心我国概率统计学科的健康发展，为了推动理论联系实际的工作，他曾建议创设统计实验室．为了促进学术交流，特别是给青年人提供发表文章的机会，他还发起筹办过全国性的概率统计杂志．由于众所周知的原因，他的这些设想生前未能实现．今天可以告慰于许宝騄先生的是，他的建议现在都已实现了：北京大学统计实验室已于 1983 年建立；中国现场统计研究会主办的《数理统计与管理》已在 1982 年创刊；中国概率统计学会主办的《应用概率统计》也在 1985 年创刊．

许宝騄逝世时享年仅 60．但是，他留下了举世瞩目的学术成就，留下了为我国概率统计学科的建立和发展所开创的业绩，也留下了热爱祖国、献身科学的宝贵精神财富．1980 年和 1990 年，在他诞辰 70 周年和 80 周年的时候，北京大学分别举行了全国性的纪念许宝騄学术讨论会，表达对他的深切怀念．今年，我们纪念他诞辰 90 周年，要进一步学习他的奋斗精神，为创建世界一流的北京大学，为使我国的数学学科早日达到世界先进水平而不懈努力．

习题参考答案

12.1.1 练习

$\bar{y}=\frac{1}{n}(y_1+y_2+\cdots+y_n)=\frac{1}{n}(x_1+b+x_2+b+\cdots+x_n+b)=\frac{1}{n}(x_1+x_2+\cdots+x_n)+b=\bar{x}+b.$

习题 1

1. 略

2. 证明：$\bar{y}=\frac{1}{n}$（$y_1+y_2+\cdots+y_n$）$=\frac{1}{n}$（$ax_1+ax_2+\cdots+ax_n$）$=a\cdot\frac{1}{n}$（$x_1+x_2+\cdots+x_n$）$=a\bar{x}$.

3. （1）$\mu_1=60.25$，$\mu_2=63.625$，$\mu_3=60.625$，$\mu=61.5$.

（2）有. $\because(\mu_1+\mu_2+\mu_3)/3=(\frac{1}{8}\sum_{i=1}^{8}X_i+\frac{1}{8}\sum_{i=9}^{16}X_i+\frac{1}{8}\sum_{i=17}^{24}X_i)/3=\frac{1}{24}\sum_{i=1}^{24}X_i$，$\therefore(\mu_1+\mu_2+\mu_3)/3=\mu$.

12.1.2 练习

证明：$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i=\frac{1}{n}\sum_{i=1}^{n}(ax_i+b)=a\cdot\frac{1}{n}\sum_{i=1}^{n}x_i+b=a\bar{x}+b.$

习题 2

1. 总体均值是总体的指标，是一个固定的量，而样本均值带有随机性，随着样本 n 的增加，样本均值 $\bar{x}$ 会接近 μ.

2. 略

3. $\bar{x}=688.65$

4. 相等. 设 a_{11}，a_{12}，…，a_{15}，a_{21}，a_{22}，…，a_{25}，…，a_{n1}，a_{n2}，…，a_{n5} 为 $5n$ 个个体，由题意得，

$\mu_1=\frac{1}{5}$（$a_{11}+a_{12}+\cdots+a_{15}$），$\mu_2=\frac{1}{5}$（$a_{21}+a_{22}+\cdots+a_{25}$），…，$\mu_n=\frac{1}{5}$（$a_{n1}+a_{n2}+\cdots+a_{n5}$），

于是，

$\frac{1}{n}$（$\mu_1+\mu_2+\cdots+\mu_n$）$=\frac{1}{5n}$（$a_{11}+a_{12}+\cdots+a_{15}+a_{21}+a_{22}+\cdots+a_{25}+\cdots+a_{n1}+a_{n2}+\cdots+a_{n5}$）$=\mu$.

高中数学教师教学用书

12.1.3 练习

$\because \bar{y}=\bar{x}+b$,

$\therefore s_y^2=\frac{1}{n}\sum_{i=1}^{n} y_i^2-\bar{y}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i+b)^2-(\bar{x}+b)^2$

$=\frac{1}{n}\sum_{i=1}^{n} x_i^2+2b\bar{x}+b^2-(\bar{x}^2+2b\bar{x}+b^2)=\frac{1}{n}\sum_{i=1}^{n} x_i^2-\bar{x}^2=s_x^2$.

习题 3

1. $\bar{x}\approx194.38$，$s^2\approx16.39$，$s\approx4.05$

2. 用 $\bar{x}_{甲}$，$s_{甲}$ 和 $\bar{x}_{乙}$，$s_{乙}$ 分别表示甲，乙两厂的白糖的均值和标准差，经过计算得到，$\bar{x}_{甲}=\bar{x}_{乙}=500$，$s_{甲}<s_{乙}$，甲厂家生产的白糖质量比较稳定，销售部应当销售甲厂的白糖.

3. A 和 C

4. $\bar{x}=\frac{31}{7}$，$s^2\approx10.82$，$s\approx3.29$.

5. $\because s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i^2-\bar{x})^2=0$,

$\therefore x_i-\bar{x}=0, i=1,2,\cdots,n$,

从而所有的 x_i 都相等，且为 $\bar{x}$.

6. 记 $y_i=ax_i, s_y^2=\frac{1}{n}\sum_{i=1}^{n} y_i^2-\bar{y}^2=a^2\frac{1}{n}\sum_{i=1}^{n} x_i^2-a^2\bar{x}^2=a^2(\frac{1}{n}\sum_{i=1}^{n} x_i^2-\bar{x}^2)=a^2s_x^2$.

7. 由上题可得，$s_y^2=a^2s_x^2$，从而 $s_y=|a|s_x$.

12.2.1 练习

1. 不是　　2. 是

习题 4

1. 简单随机抽样是指无放回随机抽样，即在总体中抽出一个个体后，下次在余下的个体中再进行随机抽样的一种抽样方法.

2. 有放回随机抽样是指从总体中抽出一个个体，记录下抽到的结果后放回，摇匀后再进行下一次随机抽样的一种抽样方法.

3. 比如要调查某校全体学生的身高情况，若抽样时只抽取男生或者低年级学生进行调查，则得到的结论必然是有偏的.

4. 应该采取有放回抽样，海豚属于国家保护动物.

5. 这两个抽样方案都不行，违背了抽样随机性的原则. 前者忽略了低收入人群，后者忽视了高收入人群.

12.2.2 练习

A

习题 5

1. 略

2. 略

12.2.3 练习

（1）随机抽取大约 28 名男生和 22 名女生进行调查.

（2）第一层为男生（或女生），第二层为女生（或男生）.

（3）是的.

习题 6

1. 略

2. D

3. 总体平均值是 $\bar{x}=161.8$ cm，所选用的样本平均值为 161.6 cm.

12.3.1 练习

数据中的最小值是 150，最大值是 172，这 49 个数据就散布在闭区间 [150，172] 中，取一个略大的区间 (146，174]，将 (146，174] 七等分，排在表的第一列. 计算出数据落入各段的个数 n_i，填入第二列，计算出数据落入各段的频率，

$f_1=\frac{1}{49}\approx 2.04\%$，$f_2=\frac{1}{49}\approx 2.04\%$，$f_3=\frac{10}{49}\approx 20.41\%$，$f_4=\frac{14}{49}\approx 28.57\%$，

$f_5=\frac{17}{49}\approx 34.69\%$，$f_6=\frac{5}{49}\approx 10.2\%$，$f_7=\frac{1}{49}\approx 2.04\%$，

依次填入第三列，最后将各列之和填入最后一行，得到频率分布表.

身高	发生次数 n_i	f_i＝发生的频率
(146，150]	1	2.04%
(150，154]	1	2.04%
(154，158]	10	20.41%
(158，162]	14	28.57%
(162，166]	17	34.69%
(166，170]	5	10.2%
(170，174]	1	2.04%
总计	49	100%

从上述频率分布表可以方便地分析出以下结果：有 4.08%的同学的身高不到 154 cm，有 63.26%的同学的身高在 158～166 cm 之间，只有 2.04%的同学的身高超过 170 cm.

习题 7

（1）数据中的最小值为 259，最大值为 411，这 50 个数据就散布在闭区间［259，411］中，取一个略大的区间（250，425］，将（250，425］七等分，排在表的第一列. 计算出数据落入各段的个数 n_i，填入第二列，再计算出数据落入各段的频率，

$f_1=\frac{1}{50}=2\%$，$f_2=\frac{6}{50}=12\%$，$f_3=\frac{13}{50}=26\%$，$f_4=\frac{12}{50}=24\%$，

$f_5=\frac{7}{50}=14\%$，$f_6=\frac{7}{50}=14\%$，$f_7=\frac{4}{50}=8\%$，

依次填入第三列，最后将各列之和填入最后一行，得到频率分布表.

身高	发生次数 n_i	f_i＝发生的频率
（250，275］	1	2%
（275，300］	6	12%
（300，325］	13	26%
（325，350］	12	24%
（350，375］	7	14%
（375，400］	7	14%
（400，425］	4	8%
总计	50	100%

（2）从述频率分布表可方便地分析出以下的结果：有 14%的公交车的营业额低于 300 元，有 50%的公交车的营业额在 300～500 元之间，营业额超 400 元的公交车只占到 8%.

12.3.2　练习

略

习题 8

略

12.3.3　练习

略

习题 9

略

12.3.4　练习

数学	
6	5578
7	015679
8	02333445667888
9	133333457999
10	0000

物理	
5	7
6	02477
7	02344477889
8	00111224444678
9	3566899
10	00

语文	
6	277
7	0001223455667888899
8	000012223344689
9	035

习题 10

首先作数学和物理的双茎叶图，因为所给出的数据有两位的也有三位的，所以选用十位或百位上的数作“茎”，排在双茎叶图的中间一列，它们分别是 5，6，7，8，9，10，然后将数学的个位数作为“叶”，依次排放在相应的茎的左边，把物理数据排在“茎”的右边，如图12－14.

数学 树叶	树茎	物理 树叶
	5	7
8755	6	02477
976510	7	02344477889
88876654433320	8	00111224444678
999754333331	9	3566899
0000	10	00

图 12－14

按照上述方法类似地可作出物理和语文的双茎叶图（图 12－15）及数学和语文的双茎叶图（图 12－16）.

语文 树叶	树茎	物理 树叶
	5	7
772	6	02477
9988887665543221000	7	02344477889
986443322210000	8	00111224444678
530	9	3566899
	10	00

图 12－15

数学 树叶	树茎	语文 树叶
	5	
8755	6	
976510	7	277
88876654433320	8	0001223455667888899
999754333331	9	000012223344689
0000	10	035

图 12－16

(1) 从上述茎叶图可以看出，数学平均成绩最好.

(2) 语文成绩基本都处在 70～90 分之间，其分布最集中.

12.4.1　练习

用 x 表示年份，用 y 表示比萨斜塔倾斜量，其散点图见图 12－17.

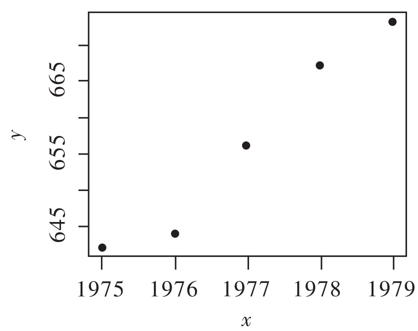


图 12－17

习题 11

用 x 表示月份，y 表示平均气温，北京和广州月平均气温的散点图分别如图 12－18 和12－19. 由图可知，气温平均值与月份无关.

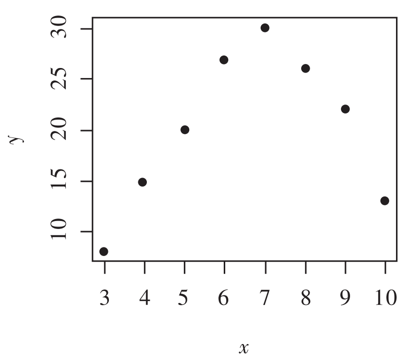


图 12－18

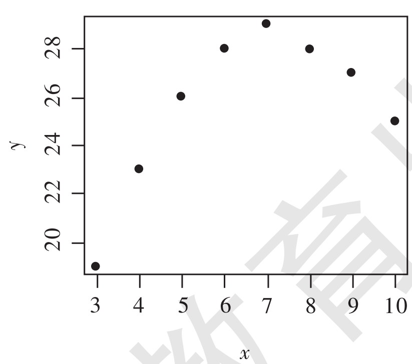


图 12－19

习题 12

1. 证明：回归直线为 $y=bx+a$，其中 $b=\frac{s_{xy}}{s_x^2}$，$a=\bar{y}-b\bar{x}$. 把 $a=\bar{y}-b\bar{x}$ 代入 $y=bx+a$ 中，我们可得 $y-\bar{y}=b(x-\bar{x})$，显然 $(\bar{y},\bar{x})$ 总在回归直线上.

2. (1) $\bar{x}=1\ 640$, $\bar{y}=15.2$, $l_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=492\ 000$,

$l_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=5\ 460$，从而 $b=\frac{l_{xy}}{l_{xx}}\approx 0.011\ 1$, $a=\bar{y}-b\bar{x}=-3$.

故回归直线为：$y=0.011\ 1x-3$.

(2) 记 $x_1=1\ 850$，$x_2=1\ 500$，把它们分别代入回归直线中，可得，

$\hat{y}_1=0.011\ 1x_1-3=17.535\approx 18$，$\hat{y}_2=0.011\ 1x_2-3=13.65\approx 14$.

(3) 如下图

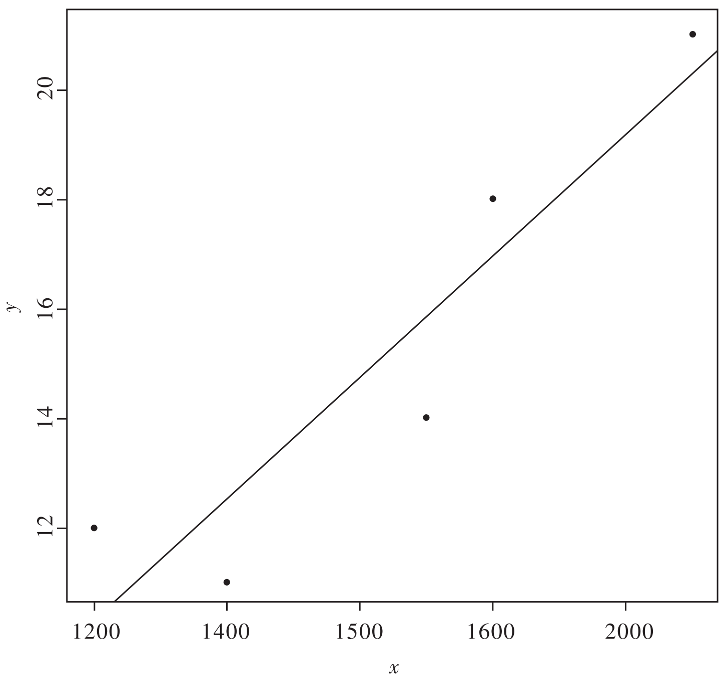


图 12 - 20

3. (1) 回归直线为：$y=-1.1153x+3.6084$；

(2) 记 $x_1=1.9$，$x_2=2.8$，$x_3=3.2$，分别代入上述回归直线，得到三个预测值分别为：$\hat{y}_1\approx 1.49$，$\hat{y}_2\approx 0.49$，$\hat{y}_3\approx 0.04$.

(3) 如下图

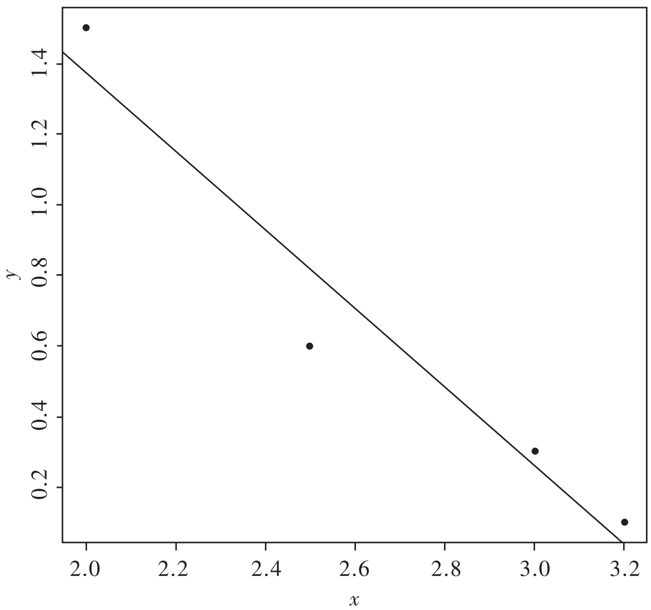


图 12－21

复习题十二

1. 都不能，原因是这样得到的样本都是有偏样本，比如只在娱乐场所抽样可能会忽视中低收入人群，而只在公共汽车站抽样可能会忽视高收入人群，这样抽样违背了抽样的随机性的原则，因而可能会导致错误的结论.

2. 样本均值为 12.25，样本标准为 7.281 54.（按计算器算出的标准差为 7.358 596）

3. 画出两人得分的茎叶图，为便于对比分析，可将茎放在中间共用，叶分居左、右两侧（如图 12－22 所示）：

甲		乙
	0	8
52	1	346
54	2	368
976611	3	389
94	4	
0	5	1

图 12－22

从图 12－22 的茎叶图可以看出，甲运动员的得分大致对称，平均得分、众数及中位数都是 30 多分．乙运动员的得分除一个 51 分外，也大致对称，平均得分、众数及中位数都是 20 多分．因此甲运动员发挥比较稳定，总体得分情况比乙好.

4. 略

5. （1）～（4）略，（5）400.3，（6）3 930

6. 略

7. 北京月平均气温的回归直线为：$y=0.112x+19.4$，广州月平均气温的回归直线为：$y=0.4547x+22.67$.

8. (1) 回归直线为：$y=0.085x+40.2$.

(2) 数据的散点图和回归直线如图 12－23 所示：

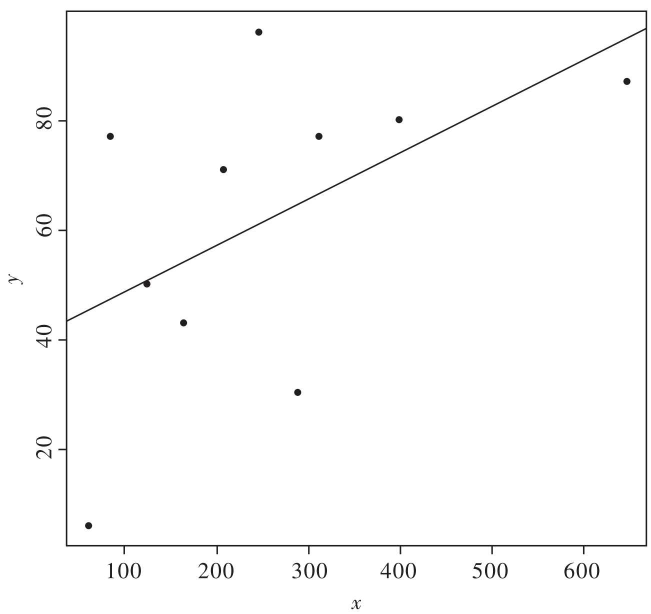


图 12－23

(3) $x=60$ 时，y 的预测值为　$y=0.085\times60+40.2=45.3$；

$x=70$ 时，y 的预测值为　$y=0.085\times70+40.2=46.2$；

$x=200$ 时，y 的预测值为　$y=0.085\times200+40.2=57.2$；

(4) $y=85$ 时，x 的预测值为　$x=(85-40.2)/0.085=527$.

9. (1) 回归直线为：$y=0.15+0.86x$.

(2) 数据的散点图和回归直线如图 12－24 所示：

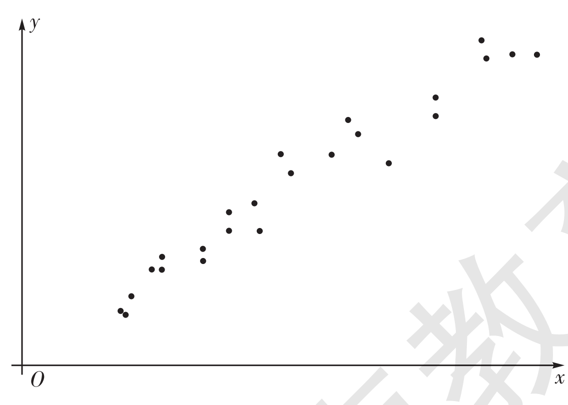


图 12－24

（3）$x=5.0$ 时，y 的预测值为：$y=0.86\times5.0+0.15=4.45$；

$x=5.2$ 时，y 的预测值为：$y=0.86\times5.2+0.15\approx4.62$.

（4）$y=4.5$ 时，x 的预测值为：$x=(4.5-0.15)/0.86\approx5.06$.

第 13 章　概率

一、教学目标

1. 在具体环境中，了解随机事件发生的不确定性与频率的稳定性，进一步了解概率的意义以及频率与概率的区别.

2. 通过实例，理解互斥事件的概率加法公式和对立事件的概率公式.

3. 通过实例，理解古典概型及其概率计算公式，会用列举法计算一些随机事件所含的基本事件数及事件发生的概率.

4. 了解随机数的意义，能运用模拟方法（包括计算器产生随机数来进行模拟）估计概率，初步体会几何概型及其概率计算公式，会计算一些简单的几何概型的事件发生的概率.

5. 通过阅读材料，了解人类认识随机现象的过程.

二、教材说明

随机现象在日常生活中随处可见，概率是研究随机现象的学科，它为人们认识客观世界提供了重要的思维方式和解决问题的方法，同时为统计学的发展提供理论基础，而且概率论的基础知识已经成为每个未来公民的必备常识. 学生应结合具体实例，学习概率的某些基本性质和简单的概率模型，加深对随机现象的理解，能通过实验，计算器（机）模拟估计简单随机事件发生的概率，消除日常生活中的一些错误认识，学会用科学的方法来观察世界和认识世界.

概率起源于机会游戏，它与第 12 章一样，都是研究随机现象或不确定现象的学科，这一点也是与纯数学相区别的地方，教师在教学中要把握这些区别.

本章首先通过实例引入频率的稳定性，然后介绍试验与事件以及事件的运算，以此作为基础，介绍古典概型及其计算，再推广到几何概率的计算，最后介绍频率与概率的关系，与引言部分遥相呼应. 本章内容首尾衔接，逻辑性强，环环紧扣，又循序渐进，符合学生的认知特点.

三、课时安排建议

本章教学约需 8 课时，具体分配如下（仅供参考）：

13.1　试验与事件

13.1.1　事件　　1 课时

13.1.2　事件的运算　1课时

13.2　概率及其计算

13.2.1　古典概率模型　2课时

13.2.2　几何概率　1课时

13.3　频率与概率

小节与复习　1课时

四、教学建议

本章的重点和难点内容为第二节概率及其计算. 古典概型和几何概型都是等可能概型，对等可能性的理解和把握是正确理解概率的比例算法的基础. 古典概型概率的计算要涉及到一些排列组合的知识，同时又需要一些技巧，古典概率的计算一直是一个难点问题，正确计算这类问题概率的关键是分析好全集和事件各自包含的元素的个数.

概率是一个与现实生活紧密相连的一个部分. 教师在教学中一定要注意教学与具体实际相结合. 使学生做到既掌握好书本知识，又能将所学知识在生活实践中予以具体应用. 既培养了学生的能力，又提高了学生学习的兴趣.

计算技术的飞速发展为数学特别是概率统计的教学提供了极大的便利，以前很多无法解决的问题现在借助计算机能轻而易举的解决. 有条件的学校可以让学生利用好计算机资源，特别是讲授频率的稳定性以及后面的蒙特卡罗试验等内容，教师可以利用Z+Z智能画板或者其他软件进行实际演习，同时引导学生也进行相应的试验，即可加深对书本知识的理解（比如频率的稳定性），又能解决一些具体问题（比如求圆周率的近似值）等，同时还可以进一步提高学生的学习兴趣和热忱.

五、评价建议

1. 本章内容在现实生活中有很多应用，可让学生留心生活中的一些现象，利用业余时间，根据本地区、本学校的实际情况，由教师指定课题或者学生自己选题，开展研究性学习，写一篇“生活中或者身边的概率问题”的小论文. 比如说针对某地下六合彩猖獗的情况，写一篇小论文，并利用自己所学的概率知识和一些实例说服身边的人不要进行地下六合彩活动.

2. 投骰子，掷硬币是概率中的一类常见问题，因此可让学生以小组为单位，将试验的过程和结果记录下来. 也可由教师提出一个问题，由学生独立完成，然后检查试验过程和结果是否合理规范，可采用教师评议、学生互评等手段.

3. 针对本章所学内容，可在学完本章知识之后来一次总的测试，以评价学生掌握书本知识的程度.

13.1 试验与事件

13.1.1 事件

教材线索

教材首先通过一个掷硬币的古老而有趣的例子引出了频率的稳定性，然后介绍试验的元素、全集，随机事件、不可能事件、必然事件等概念，同时指出一个事件是否发生的含义.

教学目标

1. 了解试验的特点，知道试验的全集、元素和事件之间的关系. 能写出试验的全集以及事件所包含的元素.

2. 了解必然事件和不可能事件的含义.

3. 能结合具体实际举出试验以及随机事件的例子.

教材分析

1. 重点

（1）对试验、随机事件的理解.

（2）试验的全集、元素和事件的关系.

2. 难点

（1）试验的全集、随机事件和基本事件的联系和区别，事件发生的含义；

（2）如何启发学生联系自身的学习和生活经历举出试验的例子.

3. 频率的稳定性

教材在引言部分介绍了频率的稳定性，为 13.3 节频率与概率作了铺垫. 所谓频率的稳定性，是指在相同的条件下，大量重复试验中出现某一事件的频率稳定在某一个固定值（概率）的附近.

关于频率的稳定性，历史上还有数人做过类似于凯瑞的试验，见本节的相关链接. 教材中的南非数学家凯瑞并非历史上最早做此试验的人.

4. 随机试验

随机试验简称试验. 教材中没有给出明确的定义和说明，但学生对这个词汇理解上应该问题不大. 一般地，试验一词有十分广泛的含义，凡是对某一现象的观察和为此进行的实验都称

之为试验．称满足以下条件的试验为随机试验，也简称为试验．

（1）试验可以在相同的条件下重复进行；

（2）每次试验可以出现不同的结果，而究竟出现哪一个结果，试验前不能预先断言；

（3）试验中一切可能的结果是事先已知的，每次试验总是恰好出现这些结果中的一个．

随机试验的每一个可能结果，称为基本事件，也称作样本点，又称为元素，常用 ω 表示．所有元素构成的集合为试验的全集，统计学家又称之为样本空间，常用 Ω 表示．

5. 事件的直观意义

教材中的事件是作为全集的子集而定义的，事件有它的直观含义．

若某件事情在一次试验中一定发生，称这件事情为必然事件．例如事件“在一个大气压下，100 ℃的水一定沸腾”．又如事件“在一副扑克牌中任摸 14 张，必有两张是异色”．显然，全集 Ω 包含了所有样本点，因而一定会发生，为必然事件．

若某件事情在一次试验中一定不发生，称这件事情为不可能事件，用符号 Φ 表示．比如事件“在一副扑克中任摸 14 张，没有两张是异色”就是不可能事件．

若某件事情在一次试验中可能发生也可能不发生，称这件事情为随机事件，简称为事件．比如事件“从一副扑克牌中任摸两张，所得的花色相同”．又如事件“从一匹含有一定次品数的产品中任意取出 3 件，其中恰好有一件次品”等等．

可见试验的每一个可能结果也是事件，因为这种事件不可能再分为更为简单的事件，所以特别称这种事件为基本事件．而一般的事件就是由若干个基本事件复合而成的．

必然事件，不可能事件不是事件（即不随机），但是为了讨论问题的方便，也把它们算作事件．

要注意事件与基本事件这两个概念的区别．基本事件为在样本空间里面不可再分的最小元素，而一个事件可以由若干基本事件组成．比如投掷一枚六个面的骰子，在该试验中出现“奇数点”的结果为一事件，但它不是基本事件，它是由三个基本事件“1 点”、“3 点”和“5 点”构成的．

6. 随机事件的集合论意义

随机事件简称事件，它是由若干个个基本事件复合而成．而基本事件就是样本点，所以事件是由若干个样本点组成的，故事件可看作是样本空间 Ω 的一个子集．所谓在一次试验中事件 A 发生，当且仅当试验中出现的样本点 $\omega\in A$；所谓在一次试验中事件 A 不发生，当且仅当在试验中出现的样本点都不属于 A．

有时全集的元素不一定是可数的，请看下例：

例．向某一目标射击一发炮弹，观察落点与目标的偏差，样本点可以是任何一个非负实数（偏差值），样本空间 $\Omega=\{d \mid d\geqslant 0\}$．事件 A：“偏差不超过 100 米”，则 $A=\{d \mid 0\leqslant d\leqslant 100\}$．

教学建议

本小节是概率论的起始节．概率与统计一样都是研究不确定现象的学科，都是与身边的实

际紧密相连的. 教师可以通过本节的相关链接举例说明之.

建议教师在讲授试验、随机事件、必然事件以及不可能事件时，尽量启发学生联系自身的学习和生活经历举出例子，以增进理解和掌握.

要多通过实例使学生学会写出试验的全集以及事件所包含的元素，为后面概率的计算打下基础.

例题分析

例 1. 将一枚硬币投两次，观察正反面出现的情况，用 H 表示正面朝上，T 表示反面朝上. 写出试验的全集和元素，并写出以下事件：

A：两次出现的面不同，B：第一次出现反面.

分析：这道例题的第一部分实际上就是教材的例 1 与例 2.

注意：这里将一枚硬币投掷两次是一次试验而不是两次试验！

教材为了防止学生混淆，特地分两道例题来说明此事. 所以，这里的全集为 $\Omega=\{HH, HT, TH, TT\}$，元素有四个，分别为 HH，HT，TH 和 TT.

事件 A“两次出现的面不同”，$A=\{HT, TH\}$；

事件 B“第一次出现反面”，$B=\{TH, TT\}$.

例 2. 从包含两件次品（记作 a_1，a_2）和三件正品（记作 b_1，b_2，b_3）的五件产品中，任意抽出两件. 请写出试验的全集. 并写出以下事件：

（1）A：没有抽到次品；（2）B：恰好抽到一件次品；

（3）C：至少抽到一件次品；（4）D：至多抽到一件次品.

解：每抽出两件产品就是一次试验，例如拿出的两件是 a_1 和 b_3，这就是一个元素，记作 (a_1, b_3). 所有的元素共有 10 个，故试验的全集为

$$\Omega=\{(a_1, a_2), (a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), (a_2, b_2), (a_2, b_3), (b_1, b_2), (b_1, b_3), (b_2, b_3)\}$$

考虑事件 A，在一次试验中，事件 A 发生当且仅当在这次试验中出现样本点 (b_1, b_2)，(b_1, b_3)，(b_2, b_3) 中的一个. 这样我们可以认为事件 A 是由 (b_1, b_2)，(b_1, b_3)，(b_2, b_3) 组成的集合，故

$$A=\{(b_1, b_2), (b_1, b_3), (b_2, b_3)\}.$$

同理，事件 $B=\{(a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), (a_2, b_2), (a_2, b_3)\}$.

事件 C“至少抽到一件次品”，表示抽到的两件产品中，有一件或者两件是次品. 故

$C=\{(a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), (a_2, b_2), (a_2, b_3), (a_1, a_2)\}$.

事件 D“至多抽到一件次品”，表示抽到的两件产品中两件都是正品或者只有一件是次品. 故

$D=\{(b_1, b_2), (b_1, b_3), (b_2, b_3), (a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), (a_2, b_2), (a_2, b_3)\}$.

相关链接

频率的稳定性

在投掷一枚均匀硬币时，既可能出现正面向上，又可能出现反面向上，预先做出确定的判断是不可能的. 但是，由于硬币是均匀的，直观上来看发生正面和发生反面的机会因该相等，也即在大量的试验中，发生正面的频率应该稳定在 0.5 左右，历史上曾有不少人做过试验，其结果见教材表 13.1.

下面我们再举几个例子说明频率的稳定性.

（一）一个孕妇生男还是生女是偶然的，但就整个国家和大城市来说，从人口普查的资料中可以看出，男孩出生数占全体出生数的比例几乎是年年保持不变的，这反映男孩出生频率的稳定性. 在古代中国，公元前 228 年，根据人口普查算出男孩出生的频率约等于 0.5. 十八世纪，法国数学家拉普拉斯对伦敦、彼得堡、柏林和整个法国的广大人口普查资料进行了研究，得出那些地区的男孩出生的频率约等于 22/43.

（二）曾有人统计过某个国家每年因没有写清地址或其它原因而无法投递的信件数，这从常识上来看，似乎没有什么规律性，但是经过统计后惊人地发现，一年中这类信件数在全体信件中所占有的比例许多年几乎保持不变. 如下表所示：

表 13－1

年　份	信件总数 n	无法投递的信件数 μ_n	频率 μ_n/n
1906	983×10^6	54 861	5.6×10^{-5}
1907	$1\,070\times10^6$	53 500	5.0×10^{-5}
1908	$1\,214\times10^6$	59 627	4.9×10^{-5}
1909	$1\,357\times10^6$	62 088	4.6×10^{-5}
1910	$1\,507\times10^6$	76 614	5.1×10^{-5}

这类例子不胜枚举. 这一切表明，在大量试验中事件 A 具有频率的稳定性，也就是统计规律性.

13.1.2　事件的运算

教材线索

教材通过几个具体例子，介绍了事件的交、并、补等运算. 并介绍了对立事件和互斥事件

的概念.

教学目标

1. 理解事件的交、并、补等运算，会进行事件的交、并和补等运算.

2. 理解对立事件和互斥事件的含义，能在实例中加以区分.

教材分析

1. 重点

事件的并、交和补运算. 互斥事件和对立事件的概念.

2. 难点

如何确定每个事件所包含的元素. 怎样在实例中区分互斥事件和对立事件.

3. 本小节主要讲授事件的各种运算，包括并、交和补. 本小节的内容主要是为下节概率及其计算作铺垫.

4. 事件就是集合，因此对事件可以进行并、交和补的运算. 对事件进行各种运算，最关键的是确定全集和事件所包含的元素. 正确分析和计算全集和事件所包含的基本事件个数，是古典概型的概率计算基础.

5. 事件的交、并和补

教材只通过例题介绍了事件的交、并和补等运算，但没有给出它们的直观含义或者定义. 下面用概率论的语言给出它们的含义.

事件 A 和事件 B 的交，是指事件 A 与事件 B 同时发生，记作 $A\cap B$ 或者 AB；

事件 A 和事件 B 的并，是指事件 A 与事件 B 至少有一个发生，记作 $A\cup B$；

事件 A 和事件 B 的差，是指事件 A 发生而事件 B 不发生，记作 $A\setminus B$；

若事件 A 和事件 B 不能同时发生，则说明事件 A 和事件 B 互斥；

记 $\bar{A}=\Omega\setminus A$，则称事件 $\bar{A}$ 为事件 A 的对立事件.

由于事件间的关系和运算与集合论的对应关系如下：

概率论	集合论
事件空间	全集 $\Omega=\{\omega\}$
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	$A\setminus B$
事件 A 与 B 互斥	$A\cap B=\Phi$
事件 A 与 B 对立	$A\cap B=\Phi$ 且 $A\cup B=\Omega$

6. 对立事件和互斥事件

教师可通过掷骰子和抛硬币等例子来引入互斥事件和对立事件的概念. 这两个概念很重要，它们是以后讲授概率的加法公式和概率计算的一个基础.

怎样理解互斥事件和对立事件呢？互斥事件在一次试验中不可能同时发生. 两个对立事件在一次试验中非此即彼，不可能再出现其他情况.

对立事件为互斥事件，但互斥事件一般不为对立事件，只有这两个互斥事件的并为全集时，它们才为对立事件. 设 Ω 为全集，则 $B=\Omega\setminus A$ 为 A 的对立事件，又称 B 为 A 的补集. 显然，$A\cap B=\Phi$，从而它们为互斥事件.

例 1. 投掷一枚骰子，观察其出现的点数.

记事件 A：奇数点朝上；事件 B：偶数点朝上；事件 C：4 点朝上.

由事件 $A\cup B=\Omega$ 且 $A\cap B=\Phi$ 知，事件 A 和事件 B 互为对立事件，当然也是互斥事件. 事件 A 和事件 C 为互斥事件，但不为对立事件. 所有的基本事件都是互斥事件.

教材中出现了 $A\setminus B$，但没有对其进行说明. $A\setminus B$ 表示 A 与 B 的差，即在一次试验中事件 A 发生但事件 B 不发生. 用数学语言描述就是：若 $x\in A\setminus B$，则 $x\in A$ 且 $x\notin B$. 从这里可以看出，集合在计算集合 A 和 B 的差时，B 不必包含于 A.

教学建议

1. 教师在教学中可以借助集合论的文氏图来直观描述各种事件的运算关系.

2. 对于全集中的元素为有限时，常用列举法将全集以及每个事件所包含的基本事件列举出来. 这样既有利于对事件进行各种运算，又为后面学习概率的计算打下基础.

例题分析

例 2. 铅笔盒中有圆珠笔 3 支，钢笔 3 支. 从中无放回地任取 3 支，用集合 A，B，C，D，E 表示下面（1），（2），（3），（4），（5）中的事件.

（1）3 支都是圆珠笔；

（2）恰有 2 支圆珠笔；

（3）恰有 1 支圆珠笔；

（4）3 支都是钢笔；

（5）至少有 1 支圆珠笔；

讨论 A，B，C，D，E 的关系；若用它们来表示 Ω，有几种表示法？解释事件 $A\cup B$，$A\cap B$，$A\setminus B$，$\Omega\setminus A$ 的含义.

分析：本题是教材 P. 119 的例 2 的拓广.

解：采用与教材例 2 相同的记号，我们有：

（1）$A=\{y_1y_2y_3\}$；

（2）$B=\{y_1y_2g_1, y_1y_2g_2, y_1y_2g_3, y_1y_3g_1, y_1y_3g_2, y_1y_3g_3, y_2y_3g_1, y_2y_3g_2, y_2y_3g_3\}$；

（3）$C=\{y_1g_1g_2, y_2g_1g_2, y_3g_1g_2, y_1g_1g_3, y_2g_1g_3, y_3g_1g_3, y_1g_2g_3, y_2g_2g_3, y_3g_2g_3\}$；

（4）$D=\{g_1g_2g_3\}$；

（5）$E=\{y_1y_2y_3, y_1y_2g_1, y_1y_2g_2, y_1y_2g_3, y_1y_3g_1, y_1y_3g_2, y_1y_3g_3, y_2y_3g_1, y_2y_3g_2, y_2y_3g_3, y_1g_1g_2, y_2g_1g_2, y_3g_1g_2, y_1g_1g_3, y_2g_1g_3, y_3g_1g_3, y_1g_2g_3, y_2g_2g_3, y_3g_2g_3\}$

（6）由 $A\cap B=B\cap C=C\cap D=D\cap A=\Phi$ 知，事件 A，B，C，D 彼此互为互斥事件；事件 $E=A\cup B\cup C$；事件 D 和事件 E 互为对立事件，即 $D\cap E=\Phi$，且 $D\cup E=\Omega$，此也是表示全集的一种方法. 此外，还有 $A\cup B\cup C\cup E=\Omega$，$A\cup B\cup C\cup D\cup E=\Omega$. 可见全集的表示法不唯一.

注：E 中包含了“至少”这个词，一般来说，这类问题转化为求其对立事件比较容易. 此例中，$E=\Omega\setminus D$，显得比较简单，特别是在后面概率的计算时，尤为方便.

（7）$A\cup B=$“至少有 2 支圆珠笔”；$A\cap B=\Phi=$“不可能事件”；$A\setminus B=A=$“三支都是圆珠笔”；$\Omega\setminus A=$“至少有一支钢笔”.

相关链接

概率论的起源

概率论起源于机会游戏. 它的某些思想在公元前 220 年就已出现于中国的文献里. 不过它的真正历史被公认为从 17 世纪中叶开始. 1654 年法国有个叫 De Mere 的赌徒向数学家 Pascal（1623－1662）提出一个如何分赌注的问题（也叫分点问题）：简略地说，就是甲、乙两个赌徒下了赌注就按某种方式赌了起来，规定甲胜一局甲就得一分，乙胜一局乙也得一分，且谁先得到某个确定的分数谁就赢得所有赌注. 但是，在谁也没有获得确定的分数之前赌博就因故中止了. 如果甲需得 n 分才能获得所有赌注，乙需得 m 分才能获得所有赌注，问该如何分这些赌注呢？为解决这一问题，Pascal 与当时享有很高声誉的数学家 Fermat（1601－1665）建立了联系，从而使很多有名的数学家对这一问题产生了浓厚的兴趣，并使得概率论这个新领域得到了迅速的发展. 在概率论的发展史上，分赌注问题是一个极其著名的问题. 有些人把 Pascal 与 Fermat 建立联系的日子（1654 年 7 月 29 日）作为概率论的生日.

对概率论的发展做出杰出贡献的还有：17～18 世纪的惠更斯（Huygens）、贝努里（Bernoulli）、德莫佛尔（De Moivre）、辛普生（Simpson）、蒲丰（Buffon）；19 世纪的拉普拉斯

（Laplace）、高斯（Gauss）、泊松（Poisson）、契比雪夫（Chebyshev）、马尔可夫（Markov）；20世纪的柯尔莫哥洛夫（Kolmogorov）、辛钦（Khintchine）、杜布（Doob）、莱维（Levy）等等．我国现代著名的数学家侯振挺教授也因他的“侯氏定理”获得1978年英国戴维逊奖而享誉国际概率学界，他的博士生邹捷中教授也获得了1987年英国戴维逊奖．中国科学院院士，中国科学院数学与系统科学研究院马志明研究员，也因为在概率论上的杰出贡献而应邀于1994年第21届在苏黎世举行的世界数学家大会上作45分钟报告，并且在2006年第24届在西班牙马德里举行的世界数学家大会上高票当选国际数学家联盟副主席！

——摘自孙荣恒：《应用概率论》，有改动

13.2 概率及其计算

13.2.1 古典概率模型

教材线索

教材首先通过一个掷硬币的例子给出古典概型以及古典概型概率的定义. 然后再通过一些例子介绍古典概型概率的计算. 然后介绍概率的几个基本性质，并通过一些实例介绍了概率的加法公式和对立事件的概率公式. 最后指出人们在日常生活中关于概率认识的误区.

教学目标

1. 会正确判断古典概型，能通过分析事件所包含的元素和全集所包含的元素个数来计算古典概型的概率.

2. 理解概率的基本性质. 理解互斥事件概率的加法公式以及对立事件的概率公式，并能运用它们进行概率计算.

3. 能运用所学过的概率知识解答身边的一些概率问题.

教材分析

1. 重点

古典概型的概念，古典概型中随机事件的概率的计算. 概率的性质以及互斥事件概率的加法公式和对立事件的概率公式的理解和运用.

2. 难点

（1）古典概型的判断，随机事件所包含的元素个数和试验的全集所包含的元素个数.

（2）概率的加法公式的运用，如何确定随机事件的对立事件以及运用对立事件概率的公式计算概率.

3. 随机事件有其偶然性的一面，即在一次试验中它可能出现也可能不出现. 但是在大量重复试验中它又呈现内在的规律性，即它出现的可能性大小是确定的，且是可以度量的. 所谓随机事件的概率，概括起来说就是用来描述随机事件出现的可能性大小的数量指标. 它是概率论中最基本的概念之一，且是逐步形成完善起来的. 如何合理地定义概率，是一个难点问题. 本小节在最简单的情形——古典概率模型（简称古典概型）下，给出了定义概率的方法. 古典概型具有以下两个基本特征：

(1) 基本事件总数有限;

(2) 每个基本事件等可能出现.

古典概型曾经概率论发展初期的主要研究对象，它在概率论中有很重要的地位，一方面，因为它比较简单，许多概念既直观又容易理解，另一方面，它又概括了许多实际问题，有很广泛的应用.

从上面可以看出，判断一个试验是否属于古典概型，抓住那两个基本特征就可以了，其中第二个基本特征等可能性可能更难把握一些. 比如说，投掷一枚骰子，若骰子的各个面都是均匀的，则该试验属于古典概型；若骰子各个面质量分布不均匀，比如说在骰子中灌铅或水银等，赌博中称之为“出老千”，此时的随机试验就不属于古典概型. 所以在教材中，投掷硬币和骰子都指出了“均匀”二字.

4. 古典概型概率的计算

教材中古典概型概率的定义实际上给出了概率的计算方法. 从定义看来，要计算随机事件发生的概率，关键在于分析全集（样本空间）和事件各自包含的元素个数. 排列组合是计算古典概率的重要工具.

这里我们先简要复习一下计算古典概率所用到的基本计数原理.

一、加法原理

设完成一件事有 m 种方式，第一种方式有 n_1 种方法，第二种方式有 n_2 种方法，……，第 m 种方式有 n_m 种方法，无论通过哪种方法都能完成这件事，则完成这件事总共有 $n_1+n_2+\cdots+n_m$ 种方法. 例如，某人要从甲地到乙地去，可以乘火车，也可以乘轮船，火车有两班，轮船有三班，乘坐不同班次的火车和轮船，共有 $3+2$ 种方法.

二、乘法原理

设完成一件事有 m 个步骤，第一个步骤有 n_1 种方法，第二个步骤有 n_2 种方法，……，第 m 个步骤有 n_m 种方法，必须通过每一步骤才能完成这件事，则完成这件事总共有 $n_1\times n_2\times\cdots\times n_m$ 种不同的方法. 例如，一个男人有三顶帽子和两件背心，他可以有 3×2 种打扮.

加法原理和乘法原理是两个很重要计数原理，它们不但可以直接解决不少具体问题，同时也是推导下面常用排列组合公式的基础.

三、排列、组合的几个简单公式

1. 排列：从 n 个不同元素取 k 个（$1\leqslant k\leqslant n$）的不同排列总数为：

$$\mathrm{P}_n^k=n(n-1)(n-2)\cdots(n-k+1)=\frac{n!}{(n-k)!}$$

$k=n$ 时称全排列，即 $\mathrm{P}_n^n=p_n=n(n-1)(n-2)\cdot\cdots\cdot 2\cdot 1=n!$.

从 n 个不同元素取 k 个（允许重复）（$1\leqslant k\leqslant n$）的不同排列总数为：$n\cdot n\cdot\cdots\cdot n=n^k$.

2. 组合：从 n 个不同元素取 k 个 $1\leqslant k\leqslant n$ 的不同组合总数为：

$$\mathrm{C}_n^k=\frac{\mathrm{P}_n^k}{k!}=\frac{n!}{(n-k)!\,k!},$$

C_n^k 常记为 $\binom{n}{k}$，称为组合系数，且有 $P_n^k = C_n^k \cdot k!$.

3. 组合系数与二项式展开式的关系

组合系数 $\binom{n}{k}$ 又常称为二项式系数，因为它出现在下面的二项式展开式中

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

利用该公式可得到许多有用的公式．令 $a=b=1$ 得

$$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n}=2^n,$$

令 $a=-1$，$b=1$ 得 $\binom{n}{0}-\binom{n}{1}+\binom{n}{2}-\cdots+(-1)^n\binom{n}{n}=0$.

由 $(1+x)^{m+n}=(1+x)^m(1+x)^n$，运用二项式展开有

$$\sum_{j=0}^{m+n}\binom{m+n}{j}x^j=\sum_{j_1=0}^{m}\binom{m}{j_1}x^{j_1}\sum_{j_2=0}^{n}\binom{n}{j_2}x^{j_2},$$

比较两边 x^k 的系数，有

$$\binom{m+n}{k}=\sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}.$$

4. 从定义，我们可以很容易得到概率的三条基本性质，其中前两条是本质的，即

(1) 非负性：$0 \leqslant P(A) \leqslant 1$；

(2) 规范性：$P(\Omega)=1$.

事实上，我们应该把可加性列入第三条基本性质，即

(3) 可加性：若 $A \cap B=\Phi$，则 $P(A \cup B)=P(A)+P(B)$.

教材中的第三条性质 $P(\Phi)=0$ 可以由上面第二和第三条基本性质推出．推导如下：

因为 $\Omega=\Omega \cup \Phi$，且 $\Omega \cap \Phi=\Phi$，从而由第二条性质 $P(\Omega)=1$ 以及第三条性质，有 $P(\Omega)=P(\Omega \cup \Phi)=P(\Omega)+P(\Phi)$，于是，$P(\Phi)=0$．所以，教材将 $P(\Phi)=0$ 列为第三条基本性质是不太合理的，但是，限于中学生的理解水平，这样处理也未尝不可.

在一些场合，古典概型概率的定义的推广会遇到一些本质性的困难，但是概率的这三条基本性质依然存在．前苏联著名数学家柯尔莫哥洛夫就是从前两条基本性质和推广的第三条性质（可列可加性）而得到概率的公理化定义，从而使概率论上升为一门严谨的数学学科！

5. 互斥事件的概率的加法公式，可以很容易地从古典概型概率的定义中利用互斥事件的性质推导出来，即：

如果 Ω 的事件 A、B 互斥，则 $P(A \cup B)=P(A)+P(B)$．可以引导学有余力的学生思考，一般情况，事件 A、B 并不互斥，应该怎样计算 $P(A \cup B)$ 呢？

显然，事件 $A \cup B$ 的元素为事件 A 的元素再加上事件 B 的元素再减去事件 A 和 B 的公共部

分的元素（从文氏图中可以直观的看出来）. 因此

$$P(A\cup B)=P(A)+P(B)-P(A\cap B).$$

6. 对立事件的概率公式实际上给出了计算概率的另一个方法. 有时一个随机事件的概率不好直接计算时，我们就可以通过求其对立事件的概率，通过对立事件的概率计算公式求出原事件的概率. 这是一种逆向思维，在教学中应该多培养学生的这种思维习惯.

一般来说，比如问题中出现“至少”等字眼时，求其对立事件的概率比较方便. 在实际计算中，是否采用计算对立事件的概率，要根据具体问题具体分析，不可一概而论.

7. 事件的独立性与独立事件的概率公式

教材并没有涉及到这一内容，但是有的习题中须用到此知识方能解出，故在此予以介绍.

（1）首先介绍事件的独立性，准确描述事件的独立性需要用到条件概率的知识. 直观的讲，一个事件的发生对另一个事件的发生的可能性大小没有影响，这时称这两个事件是彼此独立的. 在应用上，我们可以从实际背景中判断出随机事件是否独立. 比如，射击时，命中第一发与命中第二发独立，甲命中与乙命中独立；在从无限产品中抽样或者从有限产品中放回抽样时，第一次抽到废品与第二次抽到废品独立；甲患近视与乙患近视独立等等. 而从有限产品中不放回抽样时，第一次抽到废品与第二次抽到废品不独立；甲患肝炎与同桌乙患肝炎不独立.

（2）独立事件的一个简单性质：

事件 A 与 B 独立$\Leftrightarrow A$ 与 $\bar{B}$ 独立$\Leftrightarrow\bar{A}$ 与 B 独立$\Leftrightarrow\bar{A}$ 与 $\bar{B}$ 独立.

（3）独立事件的概率计算

若事件 A 与 B 独立，则 $P(A\cap B)=P(A)P(B)$，显然，我们可以把它推广到多个独立事件的情形.

例. 从中奖率为 p 的奖券中随机抽取两张，记第一张中奖的事件为 A，第二张中奖的事件为 B，显然，事件 A 和 B 是相互独立的，且 $P(A)=P(B)=p$，$P(\bar{A})=P(\bar{B})=1-p$，下面我们计算事件 $A\cup B$ 的概率. $A\cup B$ 表示随机抽取两张奖券而中奖的事件.

方法一：直接计算

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)=P(A)+P(B)-P(A)P(B)=2p-p^2.$$

方法二：利用对立事件的概率公式计算

$\overline{A\cup B}=\bar{A}\cap\bar{B}$，表示两次抽奖都未中奖，由独立事件的性质，我们有：

$$P(\bar{A}\cap\bar{B})=P(\bar{A})P(\bar{B})=(1-p)^2,$$

从而

$$P(A\cup B)=1-P(\overline{A\cup B})=1-(1-p)^2=2p-p^2.$$

本例可以推广到摸 n 张奖券而中奖的情形，此时直接计算已是太繁琐了，而用对立事件的概率公式计算则比较简便. 记事件 C：抽奖 n 次而中奖，则

$$P(C)=1-P(\bar{C})=1-(1-p)^n.$$

教材中最后提到，“当彩票的中奖率为 1/100 时，你购买 100 张彩票中奖的概率是严格小于

1 的（实际上只有 0.634）”.

我们将 $p=0.01$，$n=100$ 代入上式中，有 $P(C)=1-(1-0.01)^{100}=0.6339677$. 在本节的例 4 中我们将进一步讨论这个问题.

教学建议

本小节的内容为该章的重点之一，也是与日常生活联系非常紧密的一部分. 建议教师在教学中与生活实际结合起来. 通过概率论的学习，培养学生分析和解决实际问题的能力，排除日常生活中一些有关概率论方面的认识误区.

根据本小节内容的特点，教师可以采用“问题探究式”教学方法进行教学. 通过问题导入、问题探究、问题解决和问题评价等教学过程，学习和挖掘古典概型的概念及相关解法. 同时，在教学过程中，注意引导学生开展小组合作学习，通过举出大量的有关古典概型的实例，调动学生的学习积极性，加强互动交流，让每一个学生都充分地参与到学习活动中来.

例题分析

例 1. 从 1，2，…，9 这九个数中，无放回地随机抽取 3 个数，则这 3 个数的和为偶数的概率是（　　）

A. $\frac{5}{9}$　　B. $\frac{4}{9}$　　C. $\frac{11}{21}$　　D. $\frac{10}{21}$

分析： 记事件 A：从 1，2，…，9 这九个数中，随机抽取 3 个不同的数，则这 3 个数的和为偶数. 从 1，2，…，9 这九个数中，随机抽取 3 个不同的数，由于不考虑顺序，所以全集中元素个数为 C_9^3；而要使取出的三个数的和为偶数，有两种方式，第一种方式从 1，3，5，7，9 中抽取两个，在 2，4，6，8 中抽取一个，包含的元素个数为 $C_5^2\cdot C_4^1$，第二种方式从 2，4，6，8 中抽取 3 个，所包含的元素个数为 C_4^3，所以事件 A 从而所求的概率

$$P(A)=\frac{C_5^2\cdot C_4^1+C_4^3}{C_9^3}=\frac{11}{21},$$

故应选答案 C.

例 2. 在一个小组中有 8 名女同学和 4 名男同学，从中任意地挑选 2 名同学担任交通安全宣传志愿者，那么选到的两名都是女同学的概率是________（结果用分数表示）.

分析： 记事件 A：选到的两名都是女同学. 从 12 名同学中任意挑选 2 名同学，由于不考虑顺序，因而有 $C_{12}^2=66$ 种选法，因此全集包含 66 个元素，而事件 A 包含的元素个数为 $C_8^2=28$，因此所求概率为 14/33.

例 3. 盒中装着标有数字 1，2，3，4 的卡片各 2 张，从盒中任意任取 3 张，每张卡片被抽出的可能性都相等，求：

（Ⅰ）抽出的 3 张卡片上最大的数字是 4 的概率；

（Ⅱ）抽出的 3 张中有 2 张卡片上的数字是 3 的概率；

(Ⅲ) 抽出的3张卡片上的数字互不相同的概率.

解: 分别用集合A,B,C表示以上(Ⅰ),(Ⅱ),(Ⅲ)中的事件. 由题意得,从盒中任意取3张的选法有C_8^3种,事件A的对立事件$\overline{A}$为抽出的3张卡片上最大的数字不是4,则$\overline{A}$的可能取法有C_6^3种,事件B的取法有C_6^1种,事件C取法有$C_4^3 \cdot C_2^1 \cdot C_2^1 \cdot C_2^1$种(把相同数字的卡片看成一对,总共有4对,第一步从4对中任取3对,有C_4^3种取法,第二步从取出的3对中,每一对任取一只的取法有$C_2^1 \cdot C_2^1 \cdot C_2^1$种),所以$P(A)=1-\dfrac{C_6^3}{C_8^3}=1-\dfrac{20}{56}=\dfrac{36}{56}=\dfrac{9}{14}$,$P(B)=\dfrac{C_6^1}{C_8^3}=\dfrac{6}{56}=\dfrac{3}{28}$,$P(C)=\dfrac{C_4^3 \cdot C_2^1 \cdot C_2^1 \cdot C_2^1}{C_8^3}=\dfrac{32}{56}=\dfrac{4}{7}$.

例4. 某福利彩票,宣称一等奖的中奖概率为万分之一,请计算摸奖一万次至少中一次一等奖的概率. 若摸奖十万次又当如何?

解: 根据本题的特点,宜采用对立事件的概率计算公式进行计算. 记事件A:"摸奖10 000次,至少中一次一等奖",则它的对立事件为$\overline{A}$:"摸奖10 000次,一次奖都不中". 从而$P(A)=1-P(\overline{A})$. 下面计算$P(\overline{A})$.

由于中奖的概率为1/10 000,则在一次摸奖中未中奖的概率为9 999/10 000=0.999 9. 摸奖一万次都未中奖的概率为$P(\overline{A})=0.999\,9^{10\,000}=0.367\,861\,046$. 可见$P(A)=1-P(\overline{A})=1-0.367\,861\,046=0.632\,138\,954$.

记事件B:"摸奖十万次至少中一次一等奖",采用类似的方法,我们可算出

$P(B)=1-0.999\,9^{100\,000}=0.999\,954\,623$.

从本题我们可见:

(1) 计算随机事件的概率时,若直接计算比较困难时,我们可以通过计算其对立事件的概率,然后利用对立事件的概率计算公式来算出其概率.

(2) 人们在日常生活中常常有这样的认识误区:"中奖率为万分之一的奖券,摸奖一万次中奖的概率为10 000×1/10 000=1,即必定中奖". 从这里我们可以看出,这种认识是荒谬的. 更浅显的一点的例子为:投掷一枚均匀的硬币,正面朝上的概率是二分之一,则连续投掷两次,必定出现一次正面. 显然是错误的.

(3) 一个随机事件,不管它发生的概率多么的小,但是在大量的重复试验中出现的概率还是很大的. 比如本题,摸奖十万次,则几乎必然会中奖.

注意: 本例涉及到独立事件的概率的计算问题,超出了学生的认知水平,但是教材中有这类问题的描述,故教师必须掌握此类问题的解法.

相关链接

一、高额奖金后面有陷阱

一项抽奖活动里的数学

大千世界，无奇不有．骗子的骗术，层出不穷．大街上一个搞免费抽奖的地摊，围了不少的旁观者，亦有许多的参与者．其具体方法是：在一个碗中放有20个红白各半的小玻璃珠，参与者从碗中随意抓出10个小玻璃珠，如抓到：

1.10个颜色全同，即10个全红或10个全白，可得一等奖，奖金300元；

2.9个颜色全同，即9个红色、1个白色或9个白色、1个红色，获二等奖，奖金30元；

3.8个颜色全同，即8个红色、2个白色或8个白色、2个红色，获三等奖，奖金3元；

4.7个颜色全同，即7个红色、3个白色或7个白色、3个红色，获四等奖，奖金2元；

5.6个颜色全同，即6个红色、4个白色或6个白色、4个红色，获五等奖，奖金1元；

6.5个颜色全同，即5个红色、5个白色，则罚款5元．

乍一看来，这类抽奖非常优惠：一是免费抽奖，多抽多奖；二是六种抽奖结果中竟有五种对参与者有利，只有一种对参与者不利；三是一等奖奖金高，而仅有一种不利于抽奖者的也只不过是罚款5元钱而已．于是乎，认为有利可图者众，纷纷参与，结果是大呼上当．然而尽管赔了钱，还是想不通，往往责怪自己运气不佳，手气不好，还想再抽一次．

其实，这是一个古典概型的概率计算问题．只要我们稍用排列组合的知识来剖析一下，便可发现在这类抽奖活动中，摊主是站在绝对有利的位置，包赚不赔的．这类问题，在概率论中，是典型的摸球问题，由排列组合知识，我们不难知道：从20个相同的球中，随机摸出10个球有$P_{20}^{10}=184\ 756$种可能结果，且这184 756个可能结果中每一种结果的出现的可能性均是相同的．不失一般性，设在某次抽奖中，摸到i个红球，则摸到的白救数是$(10-i)$个，其结果数目等于从10个红球中摸到i个球的组合数乘以从10个白球中摸到$(10-i)$个球的组合数，即为$P_{10}^{i}P_{10}^{10-i}$，$i=1$，…，10．

为了方便起见，我们不妨将摸奖结果列个表：

表13－2

摸奖结果	全红 或 全白	9红1白 或 9白1红	8红2白 或 8白2红	7红3白 或 7白3红	6红4白 或 6白4红	5红5白 或 5白5红	基本事件总数
组合公式	$P_{10}^{10}\cdot P_{10}^{0}$或 $P_{10}^{0}\cdot P_{10}^{10}$	$P_{10}^{9}\cdot P_{10}^{1}$或 $P_{10}^{1}\cdot P_{10}^{9}$	$P_{10}^{8}\cdot P_{10}^{2}$或 $P_{10}^{2}\cdot P_{10}^{8}$	$P_{10}^{7}\cdot P_{10}^{3}$或 $P_{10}^{3}\cdot P_{10}^{7}$	$P_{10}^{6}\cdot P_{10}^{4}$或 $P_{10}^{4}\cdot P_{10}^{6}$	$P_{10}^{5}\cdot P_{10}^{5}$	P_{20}^{10}
组合数	1＋1＝2	100＋100 ＝200	2 025＋2 025 ＝4 050	14 400＋14 400 ＝28 800	44 100＋44 100 ＝88 200	63 504	184 756
概率	0.000 01	0.001 08	0.021 92	0.155 88	0.477 39	0.343 72	1

表 13－3

奖项	一等奖	二等奖	三等奖	四等奖	五等奖	罚款
奖金	300 元	30 元	3 元	2 元	1 元	5 元
奖金总额	300×2＝600 元	30×200＝6 000 元	3×4 050＝12 150 元	2×28 800＝57 600 元	1×88 200＝88 200	5×63 504＝317 520

从表 13－3 可以看出，每抽奖 184 756 次，摊主付出 600＋6 000＋12 150＋57 600＋88 200＝164 550 元，而抽奖者（往往是游客）平均被罚：317 520 元！对比结果是摸奖人员损失：317 520－164 550＝152 970 元！平均每摸奖一次，摸奖人员损失：152 970÷184 756＝0.827 956 元！

此外，假若是一个公平的抽奖的话，则其奖金与罚金应大致相等，但这类抽奖的平均奖金为 164 550÷184 756＝0.890 634（元），平均罚金为 317 520÷184 756＝1.718 59（元）. 足见其罚金大大超过奖金！若此摊主再精明一点，将原来设为五等奖的与原来设为罚款的交换一下，则摊主盈利更丰！

从以上分析可以发现，这个抽奖条件表面看来很优惠，具有很大的欺骗性，亦很有诱惑力，就像猎人设置好陷阱，等待着猎物上钩似的. 这类游戏绝不是设地摊的凡夫俗子所能想出来的法子，其背后肯定有个懂概率统计的人. 你看，免费摸奖，且一等奖奖金高达 300 元，但其设置在概率竟是如此之低的事件上（约十万分之一）；罚金虽然不是很高，仅是 5 元钱而已，但其概率却高达 0.343 7！这是许多参加摸奖的人料想不及的，也是无法想通的，以致摸奖挨罚了，往往只自叹倒霉，怪自己时运不佳、手气不好，且还想再试一试！

又如下例，手法与上例相同.

某个经营洗涤用品的公司推出一个非常有吸引力的促销活动：

“本公司为了答谢广大顾客长期以来对本公司产品的支持与厚爱，特推出免费抽奖活动. 抽奖方式：箱子里有 20 个球，10 个 10 分和 10 个 5 分. 从箱子里摸出 10 个球，把各个球上的分数加总，按总分设置奖项如下：

一等奖：100 分，29 寸彩电一台，价值 3 000 元	七等奖：85 分，香皂一块，价值 4 元
二等奖：50 分，DVD 一台，价值 700 元	八等奖：65 分，牙膏一盒，价值 3 元
三等奖：95 分，榨汁机一台，价值 200 元	九等奖：80 分，牙刷一把，价值 2 元
四等奖：55 分，电饭煲一个，价值 100 元	十等奖：70 分，洗衣粉一小袋，价值 0.5 元
五等奖：90 分，*XX* 牌洗发露 2 瓶，价值 50 元	十一等奖：75 分，以成本价 25 元购买某品牌洗发露 1 瓶
六等奖：60 分，*XX* 牌洗发露 1 瓶，价值 25 元	

“抽奖结果共有 11 个，即 50、55、…、100 分. 其中有 10 个分数可免费得到奖品，中奖率超过 90%.”

我们的直觉似乎和商家宣传的一致，中奖率为 10/11≈91%. 但如果你站在一旁观察抽奖情况，你会发现得十一等奖的摸奖者很多，而且就算中其他免费奖项，也大多是一些价值较低的

小奖品. 难道是箱子里暗藏机关?

实际上，箱子里没有任何“机关”，而且关键的问题在于十一个奖项出现的概率不一样. 由于与前面的问题类似，在此就不详细说明了. 我们可以看出，所谓“免费抽奖”，实际上是实为商家推销产品、获取利润的手段.

——节选自《数理统计与管理》19卷3期（2000年5月）和25卷2期（2006年3月）

二、趣味生日问题

几乎没有一个数学教师不知道概率对生活的重要性，按约瑟夫·巴特勒的说法，概率是“生活的真正指南”. 在现实生活中，人们常常对某一事件发生的概率做出估计，这种估计往往是依靠直觉进行的. 如果没有真正领会概率论的原理，很有可能做出错误的判断. 下面的这个例子说明，直觉在很多情况下是不可靠的.

某个班级有55名学生，假定每个人的生日在一年中的任何一天的概率为1/365，问该班级至少有两人生日在同一天的概率有多大?

一年的天数远大于学生的人数，似乎至少有两人的生日在同一天的概率不会很大，但下面的计算结果可能会让你感到意外. 上面的生日问题可转化为分房问题：“有n个人，每人都以同样的概率$1/N$（$n\leqslant N$）被分配到N间房的任一间中，求至少有两个人在同一房间的概率.”

设$A=$“至少有两个人在同一房间”，则$\overline{A}=$“恰有n间房中各有一人”. 把某个人分配到N间房中之一去，又N种可能，把n个人分配到N房中去共有N^n种分配法，即基本事件总数为N^n. 恰好有n间房可从N间房中任意选出，有C_N^n种选法，对每一种这样的选法，n个人又有$n!$种不同的分配法，所以$\overline{A}$包含的基本事件总数为$C_N^n\cdot n!$，故

$$P(\overline{A})=\frac{n!\cdot C_N^n}{N^n},\ P(A)=1-P(\overline{A}).$$

再回到生日问题. 一年的365天相当于365间房，55个学生相当于55个客人，即$N=365$，$n=55$. 设$B=$“至少有两人生日在同一天”，则

$$P(B)=1-P(\overline{B})=1-\frac{n!\cdot C_N^n}{N^n}=1-\frac{55!\cdot C_{365}^{55}}{365^{55}}\approx 0.99.$$

几乎必有两人生日在同一天，这与直觉有很大的差异. 更一般地，对人数为n的班级，至少有两人在同一天过生日的概率如下表所示：

表13-4

n	10	20	22	23	30	40	50	55
$P(\overline{B})$	0.88	0.59	0.52	0.49	0.29	0.11	0.03	0.01
$P(B)$	0.12	0.41	0.48	0.51	0.71	0.89	0.97	0.99

从表13-4可以看出，22人左右可望有两人同生日的概率约为0.5；50人左右几乎必有两人同生日，这是一件很有意义的事情.

13.2.2 几何概率

教材线索

教材通过实例引出几何概率的两个定义，并总结了几何概率的基本性质，最后再通过实例介绍怎样计算几何概率.

教学目标

1. 理解几何概率的两个定义，它们分别适用于不同的范围. 体会几何概率定义的合理性.
2. 会计算一些简单的几何概率.

教材分析

1. 重点

几何概率的定义和计算.

2. 难点

几何概率定义合理性的理解以及几何概率的计算.

3. 古典概型描述了试验的全集中元素总数有限且各个元素等可能出现的情况，但是对于全集中元素的总数无限而每个元素等可能出现的试验，古典概率的定义就不适用了. 为了克服这个局限性，有必要将古典概率的定义加以推广，使得推广后的定义能适应上述情形的试验，这就是几何概率. 教材中所描述的概率模型称为几何概率模型，简称几何概型（教材中没有提及这个名称）. 它具有下列两个特征：

（1）每次试验的可能结果有无限多个，且全体可能结果可用一个有度量的几何区域来表示；

（2）每次试验的各种结果是等可能的.

兹举几例加以说明：

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，求他等待时间少于 10 分钟的概率.

例 2. 在一片面积为 1 万平方公里的海域内有面积 50 平方公里的海面下的大陆架中贮藏着石油. 现在在这片海域内任取一点钻探，求钻到石油的概率.

例 3. 在 400 毫升水中有一大肠杆菌，今在其中随机地抽取 2 毫升用显微镜观测，求观测到大肠杆菌的概率.

在例 1 中，电台每个小时报时一次，我们自然认为该人在两次报时之间醒来，且取各点的

可能性一样．一个小时 60 分钟，要使等待时间少于 10 分钟，相应的概率为$\frac{10}{60}=\frac{1}{6}$；例 2 中，由于选点的随机性，可以认为该海域中各点选中的可能性一样．故所求概率为贮油海域面积与整个海域面积之比 50/10 000；例 3 中，由于取水样的随机性，所求概率为水样的体积与总体积之比 2/400.

4. 几何概型的一般定义

设试验 E 的全集为某个可度量的区域 Ω，且 Ω 中任一区域出现的可能性的大小与该区域的几何度量成正比而与该区域的位置和性质无关，则称试验 E 为几何概型的．如果 A 是 Ω 中的一区域，且 A 可以度量，则定义 A 的概率为：

$$P(A)=\frac{A\text{ 的几何度量}}{\Omega\text{ 的几何度量}},$$

其中，如果 Ω 是一维的、二维的、三维的，则 Ω 的几何度量分别是长度、面积和体积，这样定义的概率为几何概率．显然，教材中的定义 1 和定义 2 分别对应于这里的 Ω 为一维和二维的情形.

上面的例 1、例 2 和例 3 分别对应于 Ω 的几何度量分别是长度、面积和体积的情形.

从上述定义可知，在几何概型中，“等可能”一词应理解为对应于每个试验结果的点落入某区域内的可能性大小仅与该区域的度量成正比，而与区域的位置和形状无关.

5. 几何概型和古典概型的异同

相同点：(1) 都属于等可能概型；(2) 计算方法相同，都属于概率的比例算法.

不同点：全集中元素的个数．几何概型的全集含有无穷多个元素，而古典概型只有有限个元素.

6. 几何概率的基本性质

教材中列举了几何概率的 5 条基本性质，其中第一条非负性、第二条规范性和第四条可加性是本质的，第三条 $P(\Phi)=0$ 和第五条对立事件的概率公式可由上述三条基本性质推导出来．其中第三条 $P(\Phi)=0$ 在上小节古典概型中推导过，第五条实际上也很简单：记 $\bar{A}=\Omega\setminus A$，显然 $A\cap\bar{A}=\Phi$，$A\cup\bar{A}=\Omega$，从而由可加性：

$$P(A)+P(\bar{A})=P(\Omega)=1.$$

虽然几何概率与古典概率的定义和适用范围不同，但是这三条基本性质却是一样的．前苏联数学家柯尔莫哥洛夫正是从这三条基本性质引申出来，得到了概率的公理化定义.

教学建议

建议教师通过丰富的实例引入几何概率的教学，引导学生自主发现几何概率的特征和计算公式，比如教师可以通过本小节教材分析中的三个例子引入几何概率的教学.

引导学生比较和分析几何概型与古典概型的异同，并比较古典概率的基本性质和几何概率的基本性质，从中得出概率的一些共性来.

例题分析

例 1. 向半径为 R 的圆内任意投掷一点，求此点也落在与圆内接的确定的下列图形内的概率：（1）正方形；（2）正三角形.

解： 记事件 A 为“点落入圆内接正方形内”，事件 B 为“点落入圆内接正三角形内”. 由于点是等可能的落入到圆中的，因而试验为几何概型. 全集 Ω 的度量为圆的面积 πR^2，事件 A 的度量为圆内接正方形的面积 $2R^2$，事件 B 的度量为圆内接正三角形的面积 $\frac{3\sqrt{3}}{4}R^2$. 因此

$$P（A）=\frac{A\text{ 的几何度量}}{\Omega\text{ 的几何度量}}=\frac{2R^2}{\pi R^2}=\frac{2}{\pi},$$

$$P（B）=\frac{A\text{ 的几何度量}}{\Omega\text{ 的几何度量}}=\frac{\frac{3\sqrt{3}}{4}R^2}{\pi R^2}=\frac{3\sqrt{3}}{4\pi}.$$

例 2. 已知点 B 任意地落到长度为 1 的线段 OA 上，求线段 OA 和 BA 中较短的线段大于 1/3 的概率.

解： 试验的全集为线段 OA 的长度 1，要使 OA 和 BA 中较短的线段大于 1/3，则 B 点只能位于线段 EF 内，如图 13－1 所示. 而线段 EF 的长度为 1/3，因此所求的概率为 1/3.

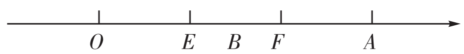


图 13－1

相关链接

贝特朗悖论

著名的贝特朗悖论是 19 世纪讨论过的几何概率的问题. 有如下表述：在单位圆内任作一弦，求其长超过圆内接正三角形边长$\sqrt{3}$的概率. 此问题可以有三种不同的解答：

解一： 先固定弦的一个端点 C 在圆周上，以 C 为顶点作圆的内接正三角形 ABC，当且仅当弦位于$\angle C$ 内部时，弦长大于$\sqrt{3}$，此时弦的另一个端点 D 必定在弧 ADB 上. 设所有方向是等可能的，因此所求的概率为弧 ADB 之长与圆周长之比，即为$\frac{1}{3}$（如图 13－2 所示）.

解二： 先固定弦的方向，由弦长与圆心距的关系可知，当且仅当弦心距小于$\frac{1}{2}$时，其弦长大于$\sqrt{3}$. 设所有交点是等可能的，因此所求概率为线段 EF 之长与直径 AB 之长的比值为$\frac{1}{2}$，其中 E、F 分别为半径 OA、OB 的中点（如图 13－3 所示）.

解三：弦被其中点位置唯一确定．只有当弦的中点 D 落在半径为$\dfrac{1}{2}$的同心圆内时，弦长大于$\sqrt{3}$．设中点位置都是等可能的，则所求概率为小圆面积与大圆面积之比$\dfrac{1}{4}$（如图 13－4 所示）．

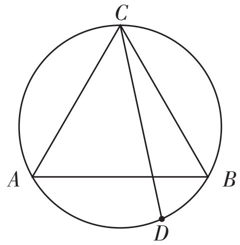


图13-2

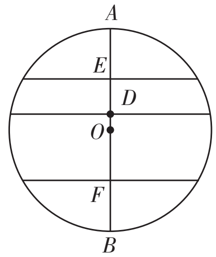


图13-3

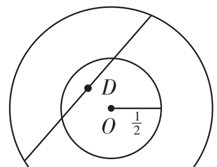


图13-4

同一问题有三种不同的答案，究竟孰是孰非呢？细究其原因，发现是在取弦时采用不同的等可能性假定．在第一种解法中，假定端点在圆周上均匀分布；在第二种解法中是假定弦的中点在直径上均匀分布；而在第三种解法中又假定弦的中点在圆内均匀分布．这三种答案是针对三种不同的随机试验．对于各自的随机试验而言，它们都是正确的．

——节选自复旦大学编：《概率论》第一册

13.3 频率与概率

教材线索

教材首先给出频率的定义，然后介绍了关于频率稳定性的一个著名例子，最后介绍计算机模拟试验及其应用.

教学目标

了解频率的稳定性，了解计算机模拟试验的原理和方法.

教材分析

1. 重点

(1) 频率的稳定性——概率的统计定义.

(2) 计算机模拟试验——蒙特卡罗方法.

2. 难点

利用蒙特卡罗方法估计未知的参数.

3. 本节主要讨论了频率和概率的关系，与本章的引言部分相应. 古典概率和几何概率的定义都要强调“等可能性”，可是一则有许多试验并非等可能的，再则，在现实生活中如何去判断“等可能”是一个很大的难题. 因此，古典概率与几何概率的定义推广到其他场合非常困难. 但是，当在相同的条件下重复多次进行同一个试验时，随着试验次数的增加，就会发现某一事件发生的频率总是稳定在某个固定的数附近，这个数就是该事件发生的概率. 这种频率的稳定性反映出了某种统计规律性，从而可以把它作为概率的统计定义.

4. 计算机模拟试验，又称为蒙特卡罗（Monte Carlo）方法，是统计学中的一个重要分支. 如今，随着计算机的飞速发展，蒙特卡罗方法已经渗入到许多领域中去了，航空航天，气象，医药卫生，工程，物理等诸多学科和领域都或多或少地要用到蒙特卡罗方法.

5. 蒙特卡罗方法

原理：根据频率的稳定性原理，采用计算机生成随机数的方法进行多次独立试验，从而确定某事件发生的概率，进而可以对未知参数进行估计.

优点：独立地重复进行某项试验，既费时，又费力，同时又损害大量财力. 特别是对于某些破坏性的试验，损失更大. 而计算机蒙特卡罗试验，则可以节省大量的时间和财力，并且可以作大量的重复模拟，省时省力高效！比如导弹在试射之前都要在计算机上作成千上万次的模拟试射.

利用蒙特卡罗方法来估计某些未知量的值，主要是利用到几何概率的等可能性原理.

教学建议

教师可以通过一些概率论发展史上一些趣味的概率问题引入本节的教学，从而可以激起学生的学习兴趣.

有条件的学校，教师可以组织学生进行计算机模拟随机试验，比如用统计软件 Matlab 或者 Z+Z 超级画板.

相关链接

一、蒙特卡罗（Monte－Carlo）方法——蒲丰的针

这里介绍一个蒙特卡罗实验的早期的一个著名的例子：1777 年法国科学家蒲丰提出下面的著名问题，这是几何概率的一个早期例子.

平面上画有距离均为 d 的一族平行线，向此平面任意投掷一长为 l（$l<d$）的针，求针与平行线相交的概率. 如图 13－5 所示.

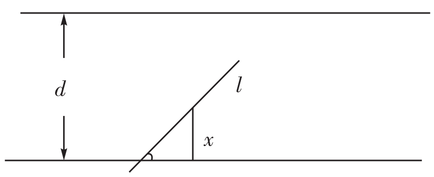


图 13－5

这个问题是一个几何概型的问题. 我们先不谈这个问题如何去求解（有兴趣的教师可以在大学概率论教材中找到解答），这个问题的重要意义在于，我们可以通过投掷针来估计 π 的值.

因为所求的概率为 $p=\frac{2l}{d\pi}$，我们只要估计出概率 p 的值来，就可以估算 π 的值了. 具体做法如下：

投针 n 次，记录针与平行线相交的次数为 μ_n，而当掷针次数很多时，由频率的稳定性，针与平行线相交的频率$\frac{\mu_n}{n}$可以作为概率$\frac{2l}{d\pi}$的近似值，因此，有

$$\pi\approx\frac{2nl}{\mu_n d}.$$

可以看到，只要投针次数 n 趋向于无穷大时，我们可以利用这种方法得到 π 的近似值. 历史上有不少学者曾利用它来计算 π 的近似值，现在把这种试验的有关资料列于下表（其中把 d 折算成 1）：

试验者	年份	针长	投掷次数	相交次数	π 的试验值
Wolf	1 850	0.8	5 000	2 532	3.159 6

Smith	1 855	0.6	3 204	1 218.5	3.154 4
De Morgan	1 860	1.0	600	382.5	3.137
Fox	1 884	0.75	1 030	489	3.159 5
Lazzerini	1 901	0.83	3 408	1 808	3.141 592 9
Reina	1 925	0.541 9	2 520	859	3.179 5

上面所用的方法概括起来就是：建立一个概率模型，使它与某些我们感兴趣的量（例如本例是 π）有关．然后设计适当的随机试验，并通过这个试验的结果来确定这些量．随着计算机的发展，已按上述思路建立了一类新的方法，即用统计实验的结果确定问题解的方法，称为蒙特卡罗（Monte－Carlo）方法，它目前已经成为统计学的一个重要分支，有着及其广泛的应用．著名的旅美华人统计学家，哈佛大学统计系刘军教授，就因为在这方面杰出的成果而获得了统计学中的最高奖—2002 年度的 COPSS 奖，他是继范剑青教授和孟晓犁教授之后，第三位来自大陆的统计学家获此殊荣！

注：COPSS 奖，一年一人，被誉为统计学中的诺贝尔奖，每年只授予一位 40 岁以下的优秀统计学家．我国吴建福（Jeff Wu，台湾），王永雄（Wing Wong，香港），范剑青，孟晓犁，刘军先后获得此项殊荣．

然而，从上表我们可以看到，这种蒙特卡罗的方法的精度确实有限．正如著名统计学家刘军教授所说的：“即使丢针丢死你也达不到祖冲之所给的精度！”我们应该为我们的先人感到自豪！

二、世界数学领袖人物——柯尔莫哥洛夫

柯尔莫哥洛夫（Andrey Nikoaevich Kolmogorov，1903－1987）出生于 1903 年，幼时失去母亲，父亲将他送到外祖父家里，由姨妈抚养长大．大约五六岁时，就发现了“$1=1\times1$，$1+3=2\times2$，$1+3+5=3\times3$，$1+3+5+7=4\times4$，…”．六岁时进入当时最好的文法学校就读．14 岁时就已涉猎高等数学．他喜欢数学，也喜欢社会科学，特别是历史．17 岁那年，有机会到莫斯科大学历史教授鲁辛（S. V. Bakhrushin）的讨论班上去听讲．柯尔莫哥洛夫回忆说：“我平生第一个科学报告是在那里做的，题目是有关 15－16 世纪诺夫哥罗德市的历史．”当时柯尔莫哥洛夫问教授能否发表这篇论文时，教授的回答是：“肯定不行，年轻人！你的证明只有一个证据，这对历史学来说，这是太少了，起码得有 5 个证据才行．”这使柯尔莫哥洛夫大失所望，以致立即喜欢起数学来：数学只要一个证明就可以了！

当然，柯尔莫哥洛夫也想过当工程师，1920 年在莫斯科大学注册时，是在冶金化学系．但不久就改变主意，改学数学，这一决定，使柯尔莫哥洛夫一辈子都没有离开过莫斯科大学的数学系．

柯尔莫哥洛夫获得许多国外科学机构的荣誉院士、荣誉会员、荣誉博士．除了荣获前苏联的国家奖、列宁奖章、契比雪夫奖、罗巴切夫斯基奖等外，还获得了一些国际奖，其中最重要

的是沃尔夫奖. 沃尔夫奖是一个终身荣誉奖，授予那些在数学领域做出个杰出贡献的数学家. 由于数学没有诺贝尔奖，因而沃尔夫奖和菲尔兹奖一起称为数学中的“诺贝尔奖”.

要概括柯尔莫哥洛夫的十分丰富的学术成就是很困难的. 当时前苏联科学院要出版他的两卷论文选集时，编者请教他如何分成两部分. 他建议，一卷是数学和力学，另一卷是概率论和信息论. 确实，20 世纪的数学有两个部分：确定性现象的数学和随机性现象的数学. 柯尔莫哥洛夫在这两个王国里面都做出了杰出的贡献.

随机现象的数学有很长的历史. 在 19 世纪 20 世纪初，它被称为“偶然性的王国”（kingdom of chance)，还没有严格的数学基础，可以说“不登大雅之堂”. 象希尔伯特、庞加莱等大数学家都没有认真地对待过它. 正是柯尔莫哥洛夫为概率奠定了严格的数学基础，对大数定律、随机变量之和等基本问题都作了十分深入的研究，特别是创立了随机过程理论，包括平稳过程，马尔可夫过程等等，打开了一片新的数学天地，并立即在通信工程、统计力学以致社会科学中获得应用，成为当今最为活跃的数学领域之一. 他的历史功绩，怎么估计都不过分.

柯尔莫哥洛夫不止在概率论方面做出了杰出的贡献，在确定性的数学方面，在纯粹数学方面，在应用数学方面都有不朽的业绩. 由于篇幅所限，在这里就不一一列举了. 他那广博的知识范围，原发性的洞察力，独特的创造性的成就，一再证明他是当之无愧的 20 世纪的数学巨人.

柯尔莫哥洛夫一直重视数学教育. 晚年从事数学数学教学改革，力图提高整个前苏联的数学教育水准. 也许因为他的数学观念太超前了，教材内容变化太大，引起各方面很大的阻力，包括来自教育部的反对和阻挠，最后终于未能成功. 这对他打击很大，以致得了帕金森症，缩短了他的生命. 1987 年，柯尔莫哥洛夫经历了辉煌的一生，却带着这份遗憾去世.

柯尔莫哥洛夫是一位伟大的教师，受他影响的数学家已不计其数. 直接受他指导的学生数目也是十分惊人，恐怕是一个难以逾越的记录：67 人，其中成为前苏联科学院院士的就有 14 人. 我国著名的概率统计学家，中国科学院院士，北京师范大学教授王梓坤先生，早年曾在前苏联求学，就是柯尔莫哥洛夫的学生.

——节选自张奠宙：《20 世纪数学经纬》

习题参考解答

13.1.1 练习

1. 试验的所有可能的结果称为元素，元素的集合称为全集.

2. 事件是全集的子集，全集也是事件，为必然事件.

3. 元素属于 A 与事件 A 发生等价.

习题 1

1. (1) $\Omega=\{i \mid i=1, 2, 3\}$，其中 i 表示取到第 i 号球，(2) $\Omega=\{(1, 2), (1, 3), (2, 3)\}$.

2. (2) $\Omega=\{(i, H), (i, T)\}$，其中 $i=1, 2, 3, 4, 5, 6$，$H=\{$正面$\}$，$T=\{$反面$\}$.

3. $A_1=\{$(正面，1)，(正面，3)，(正面，5)$\}$，
   $A_2=\{$(正面，2)，(正面，4)，(正面，6)$\}$，
   $A_3=\{$(正面，1)，(正面，2)，(正面，3)，(正面，4)，(正面，5)，(正面，6)$\}$，
   $A_4=\{$(正面，5)，(反面，5)$\}$.

13.1.2 练习

H 表示出现正面，T 表示出现反面，

$\Omega=\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$，

$A=\{HHT, HTH, THH\}$，$B=\{HHH, HHT, HTH, THH, HTT, THT, TTH\}$，

$C=\{TTT\}$，$B\setminus A=\{HHH, HTT, THT, TTH\}$，$B\setminus C=B$，

$\Omega\setminus A=\{HHH, HTT, THT, TTH, TTT\}$，$B\setminus\setminus A=$“出现三次正面或一次正面”，

$B\setminus\setminus C=B=$“至少一次正面”，$\Omega\setminus A=$“至少出现两次反面或三次正面”.

习题 2

1. (1) $A=\{HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT, THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT\}$，

$B=\{HHTT, HTHT, HTTH, HTTT, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT\}$，

$C=\{HHTT, HTHT, HTTH, THHT, THTH, TTHH, TTHT\}$，

$A\cap C=C=$“恰好 2 个反面”，$\Omega\setminus A=\{HHHH\}=$“都是正面”.

2. 用 W_i，B_j 分别表示倒出的为第 i 号白球和第 j 号黑球，$i=1, 2, 3$，$j=1, 2$，用

$W_1W_2W_3$ 表示取出的是 1 号，2 号，3 号白球，$W_1W_2B_1$ 表示取出的是 1 号，2 号白球和 1 号黑球，…，得到 $\Omega=\{W_1W_2W_3，W_1W_2B_1，W_1W_2B_2，W_1W_3B_1，W_1W_3B_2，W_2W_3B_1，W_2W_3B_2，W_1B_1B_2，W_2B_1B_2，W_3B_1B_2\}$.

（1）$A=\{W_1W_2W_3\}$；

（2）$B=\{W_1W_2W_3，W_1W_2B_1，W_1W_2B_2，W_1W_3B_1，W_1W_3B_2，W_2W_3B_1，W_2W_3B_2\}$；

（3）$C=\Omega$；

（4）$A\cup B$=“至少两个白球”$=B$，$A\cap B=A$=“3 个都是白球”，

$A\setminus B=\varnothing$=“不可能事件”，$C\setminus B=\{W_1B_1B_2，W_2B_1B_2，W_3B_1B_2\}$=“有两个黑球”.

13.2.1 练习

全集中有 $6\times6=36$ 个元素，元素（1，2）表示红骰子点数为 1，蓝骰子点数为 2，下同.

（1）$A=\{(1，1)，(2，2)，(3，3)，(4，4)，(5，5)，(6，6)\}$，$\therefore P(A)=\frac{6}{36}=\frac{1}{6}$；

（2）$B=\{(1，2)，(1，3)，(1，5)，(1，6)，(2，3)，(2，4)，(2，5)，(2，6)，(3，4)，(3，5)，(3，6)，(4，5)，(4，6)，(5，6)\}$ 共 15 个元素，$\therefore P(B)=\frac{15}{36}=\frac{5}{12}$；

（3）$C=\{(1，5)，(2，4)，(3，3)，(4，2)，(5，1)\}$ 共 5 个元素，$\therefore P(C)=\frac{5}{36}$；

（4）$\because A$ 和 B 互斥，$\therefore P(A\cup B)=P(A)+P(B)=\frac{6}{36}+\frac{15}{36}=\frac{21}{36}=\frac{7}{12}$，

又 $A\cup C=\{(1，1)，(2，2)，(3，3)，(4，4)，(5，5)，(6，6)，(1，5)，(2，4)，(4，2)，(5，1)\}$ 共 10 个元素，$\therefore P(A\cup C)=\frac{10}{36}=\frac{5}{18}$，$C\setminus B=\{(3，3)，(4，2)，(5，1)\}$，$\therefore P(C\setminus B)=\frac{3}{36}=\frac{1}{12}$.

习题 3

1. 分别以 A，B，C，D 记（1），（2），（3），（4）题中的事件.（1）$\frac{31}{365}$；（2）$\frac{62}{365}$；（3）$\frac{12}{365}$；（4）$\frac{24}{365}$.

2. 分别以 A，B，C 记（1），（2），（3）题中的事件.（1）$P(A)=0.1$；（2）$P(B)=0.9$；（3）C 即为全集，$P(C)=1$.

3. 分别以 A，B，C 记（1），（2），（3）题中的事件. 全集 Ω 为从 5 个人中录取 2 个人，一共有 10 个元素.

（1）A：一位男士和一位女士被录取，包含的元素个数为 6，$\therefore P(A)=\frac{6}{10}=\frac{3}{5}$；

（2）B：两位男士被录取，只有一个元素，$\therefore P(B)=\frac{1}{10}$；

(3) C：两位女士被录用，有 3 个元素，$\therefore P(C)=\frac{3}{10}$.

4. 分别以 A，B，C，D 记 (1)，(2)，(3)，(4) 题中的事件. 全集中有 $6\times6=36$ 个元素.

(1) A：两次骰子点数相同，包含有 6 个元素，$\therefore P(A)=\frac{6}{36}=\frac{1}{6}$；

(2) B：两次骰子点数之和为 6，包含 5 个元素，$\therefore P(B)=\frac{5}{36}$；

(3) $C=\Omega\setminus B$，$\therefore P(C)=1-P(B)=\frac{31}{36}$；

(4) D：至少有一个骰子点数为 3，则 $\overline{D}$：所得点数全不为 3，则第一枚骰子可以出现 5 个数字，第二枚骰子也可以出现 5 个数字，因此 $\overline{D}$ 中含有 25 个元素.

$\therefore P(D)=1-P(\overline{D})=1-\frac{25}{36}=\frac{11}{36}$.

5. (1) 10 种不同的结果；

(2) 摸出的 2 只球都是白球的概率为$\frac{3}{10}$；

(3) 摸出 1 只白球和 1 只黑球的概率为$\frac{3}{5}$.

6. 设 Ω 有 n 个元素，事件 A，B，C 包含的元素个数分别为 n_1，n_2，n_3，则

$P(A)=\frac{n_1}{n}$，$P(B)=\frac{n_2}{n}$，$P(C)=\frac{n_3}{n}$. 由于 A，B，C 两两互斥，所以 $A\cup B\cup C$ 中元素个数 $=A$ 中元素个数$+B$ 中元素个数$+C$ 中元素$=n_1+n_2+n_3$，于是得到

$$P(A\cup B\cup C)=\frac{n_1+n_2+n_3}{n}=\frac{n_1}{n}+\frac{n_2}{n}+\frac{n_3}{n}=P(A)+P(B)+P(C).$$

7. 由题意得，没中奖的概率为 2/3，3 次都没中的概率为$\left(1-\frac{1}{3}\right)^3=\frac{8}{27}$，则至少抽中 1 张的概率为 $1-\left(1-\frac{1}{3}\right)^3=1-\frac{8}{27}=\frac{19}{27}$.

13.2.2 练习

(1) 试验的全集是 $\Omega=[0，60]$，集合 $A=[0，40]$表示乘客候车时间多于 20 min，则$P(A)=\frac{40}{60}=\frac{2}{3}$；(2) 集合 $B=[15，15]$ 表示乘客候车时间恰好为 15 min，则 $P(A)=\frac{0}{60}=0$；(3) 集合 $C=[30，60]$表示乘客候车时间少于 30 min，则 $P(C)=\frac{30}{60}=\frac{1}{2}$；(4) 集合 $D=[25，45]$表示乘客候车时间在 25～45 min，则 $P(D)=\frac{20}{60}=\frac{1}{3}$.

习题 4

1. 记 $A=$ “麻雀落在草坪中”，$B=$ “麻雀落在花坛内”，$C=$ “麻雀落在草坪或花坛内”，

$D=$ “麻雀落在草坪或花坛外”.

(1) $P(A)=\frac{A\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{20\times 5}{400}=\frac{1}{4}$;

(2) $P(B)=\frac{B\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{\pi 5^2}{400}=\frac{\pi}{16}$;

(3) $P(C)=\frac{C\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{20\times 5+\pi 5^2}{400}=\frac{4+\pi}{16}$;

(4) $P(D)=\frac{D\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{400-20\times 5-\pi 5^2}{400}=\frac{12-\pi}{16}$.

2. 设 Ω 的面积为 1 个单位，由题意可得 A，B，C，D 的面积相等且为 $\frac{1}{4}$ 个单位，且 A，B，C，D 互不重叠，(1) $P(A)=\frac{1}{4}$；(2) $P(A\cup B)=P(A)+P(B)=\frac{1}{2}$；

(3) $P(A\cup C\cup D)=P(A)+P(C)+P(D)=\frac{3}{4}$；

(4) 因 $(A\cup B)\cap(A\cup C)=A$，$P(A\cup B)=P(A\cup C)=\frac{1}{2}$，所以有

$P((A\cup B)\cap(A\cup C))=P(A)=P(A\cup B)\cdot P(A\cup C)$.

3. 0.3.

4. 用 (x, y) 表示点的坐标，(1) 试验的全集 $\Omega=\{(x, y)\mid 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1\}$；

(2) 事件 $A=\{(x, y)\mid x^2+y^2<1, 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1\}$；(3) $P(A)=\frac{A\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{\pi}{4}$.

5. 用 (x, y) 表示点的坐标，(1) 试验的全集 $\Omega=\{(x, y)\mid 1\leqslant x\leqslant 2, 1\leqslant y\leqslant 2\}$；

(2) 事件 $B=\{(x, y)\mid |x-y|<\frac{1}{3}, 1\leqslant x\leqslant 2, 1\leqslant y\leqslant 2\}$；(3) $P(B)=\frac{B\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{5}{9}$.

6. 用 x，y 分别表示甲，乙两人到达的时刻（以小时记），

(1) 试验的全集 $\Omega=\{(x, y)\mid 1\leqslant x\leqslant 2, 1\leqslant y\leqslant 2\}$；

(2) 事件 $B=\{(x, y)\mid |x-y|\leqslant\frac{1}{3}, 1\leqslant x\leqslant 2, 1\leqslant y\leqslant 2\}$；(3) $\frac{5}{9}$.

13.3 练习

质点落入 A 的频率是 $f_N=\frac{1\,440}{3\,000}=0.48$，由频率与概率的关系可知，$f_N\approx\frac{A\text{ 的面积}}{\Omega\text{ 的面积}}$，所以 A 的面积 $\approx f_N\times\Omega$ 的面积 $=0.48\times\pi\, 12^2\approx 217\ \text{cm}^2$.

习题 5

1. (1) $\frac{50}{640}=\frac{5}{64}$;

（2）$\frac{90}{640}=\frac{9}{64}$；

（3）$1-\frac{60}{640}=\frac{29}{32}$.

2. 0.13.

复习题十三

1.（1）$\Omega=\{i \mid i=1, 2, 3, 4, 5\}$，其中 i 表示标号为 i 的球；

（2）$\Omega=\{(1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5)\}$.

2. 用 H_i 表示第 i 枚硬币正面朝上，T_i 表示第 i 枚硬币正面朝下，

$\Omega=\{(1, H_1H_2), (1, H_1T_2), (1, T_1H_2), (1, T_1T_2), (2, H_1H_2), (2, H_1T_2), (2, T_1H_2), (2, T_1T_2), (3, H_1H_2), (3, H_1T_2), (3, T_1H_2), (3, T_1T_2), (4, H_1H_2), (4, H_1T_2), (4, T_1H_2), (4, T_1T_2), (5, H_1H_2), (5, H_1T_2), (5, T_1H_2), (5, T_1T_2), (6, H_1H_2), (6, H_1T_2), (6, T_1H_2), (6, T_1T_2)\}$.

3. 用 H 表示硬币正面朝上，T 表示硬币正面朝下，试验的全集 $\Omega=\{(1, H), (1, T), (2, H), (2, T), (3, H), (3, T), (4, H), (4, T), (5, H), (5, T), (6, H), (6, T)\}$，$A=$ “硬币是正面，骰子的点数是 3”，$B=$ “硬币是正面，骰子的点数是 2 或 4”，则 $A=\{(3, H)\}$，$B=\{(2, H), (4, H)\}$.

（1）$P(A)=\frac{1}{12}$；（2）$P(B)=\frac{2}{12}=\frac{1}{6}$.

4. 由于第二子代的 D，d 基因的遗传是等可能的，而 Dd 与 Dd 的搭配方式有 4 种：DD，Dd，dD，dd，其中只有四种表现为矮茎，所以第二子代为高茎的概率为$\frac{3}{4}=75\%$.

5.（1）将骰子抛掷 1 次，它出现的点数有 1，2，3，4，5，6 这 6 种结果.

先后抛掷 2 次骰子，第一次骰子向上的点数有 6 种可能的结果，对每一种结果，第二次又都有 6 种可能的结果，于是一共有

$$6\times6=36$$

种不同的可能结果.

（2）第一次抛掷，向上的点数为 1，2，3，4，5，6 这 6 个数中的某一个，无论第一次向上的点数是多少，第二次抛掷时都可以有 2 种结果，使得两次向上的点数之和为 3 的倍数（例如，第一次向上的点数为 4，则当第二次向上的点数为 2 或 5 时，两次的点数之和都为 3 的倍数），于是共有

$$6\times 2=12$$

种不同的可能结果.

(3) 因为抛掷骰子时，出现点数 1，2，3，4，5，6 是等可能的，所以，将一枚骰子抛掷 2 次，出现的 36 个基本事件是等可能的，记“向上的点数之和是 3 的倍数”为事件 A，则事件 A 包含其中的 12 个基本事件，所以所求的概率为

$$P(A)=\frac{12}{36}=\frac{1}{3}.$$

6. (1) 事件 $A=\{(1,1)\}$，$\therefore P(A)=\frac{1}{36}$；(2) 事件 $B=\{(1,3),(3,1),(2,2)\}$，$\therefore P(B)=\frac{3}{36}=\frac{1}{12}$；(3) 事件 $C=\{(1,5),(5,1),(2,4),(4,2),(3,3)\}$，$\therefore P(C)=\frac{5}{36}$；(4) 事件 C 的概率最大，因为它所包含的元素最多.

7. (1) 0.1；(2) 0.9.

8. (1) 对任一人，其血型为 A，B，AB，O 型血的事件分别记为 A'，B'，C'，D'，它们是互斥的. 由已知，有

$$P(A')=0.28,\ P(B')=0.29,$$
$$P(C')=0.08,\ P(D')=0.35.$$

因为 B，O 型血可以输给 B 型血的人，所以“可以输出 B 型血的人”为事件 $B'+D'$. 根据互斥事件的概率加法公式，有

$$\begin{aligned}P(B'+D')&=P(B')+P(D')\\&=0.29+0.35=0.64.\end{aligned}$$

(2) 由于 A，AB 型血不能输出 B 型血的人，所以“不能输出 B 型血的人”为事件 $A'+C'$，且

$$\begin{aligned}P(A'+C')&=P(A')+P(C')\\&=0.28+0.08=0.36.\end{aligned}$$

9. (1) $\frac{4}{5}$，$\frac{19}{20}$，$\frac{22}{25}$，$\frac{23}{25}$，$\frac{89}{100}$，$\frac{91}{100}$；(2) 约为 0.9.

10. 设 Ω 有 n 个元素，事件 A，B，C，D 包含的元素个数分别为 n_1，n_2，n_3，n_4，则 $P(A)=\frac{n_1}{n}$，$P(B)=\frac{n_2}{n}$，$P(C)=\frac{n_3}{n}$，$P(D)=\frac{n_4}{n}$. 由于 A，B，C，D 两两互斥，所以 $A\cup B\cup C\cup D$ 中元素个数 $=A$ 中元素个数 $+B$ 中元素个数 $+C$ 中元素 $+D$ 中元素个数 $=n_1+n_2+n_3+n_4$，于是得到

$$P(A\cup B\cup C\cup D)=\frac{n_1+n_2+n_3+n_4}{n}=\frac{n_1}{n}+\frac{n_2}{n}+\frac{n_3}{n}+\frac{n_4}{n}=P(A)+P(B)+P(C)+P(D).$$

11. （1）$P(A)=\frac{A\text{的面积}/4}{\Omega\text{的面积}}=\frac{\pi 2^2/4}{2\times 3}=\frac{\pi}{6}$，（因为区域 A 只有 1/4 的部分含在 Ω 内）.

（2）$1-\frac{\pi}{6}$.　12. $\frac{1}{4}$.

13. 设 x，y 分别表示两人到达某地的时刻，B 表示"两人能会面"，则

（1）试验的全集 $\Omega=\{(x, y) \mid 5\leqslant x\leqslant 7, 5\leqslant y\leqslant 7\}$.

（2）两人会面当且仅当 $|x-y|\leqslant\frac{1}{2}$，故

事件 $B=\{(x, y) \mid |x-y|\leqslant\frac{1}{2}, 5\leqslant x\leqslant 7, 5\leqslant y\leqslant 7\}$.

（3）由几何概率定义，两人会面的概率为 $P(B)=\frac{7}{16}$.

14～16：由频率与概率的关系，在大量的独立试验中，频率总会在概率的左右波动，并且随着试验次数的增加，波动的幅度会越来越小．14 题中，虽然投资失败的概率为 5%，投资的次数为 n 时，失败的次数 μ_n，由频率与概率的关系，$\frac{\mu_n}{n}\approx 0.05$，因此投资失败的次数大约为 $0.05n$，显然，随着 n 的增大，必然有投资失败的时候．15 题和 16 题同样如此.

上述三题中的事件都为小概率事件，比如 14 题中投资失败的事件，15 题中射击脱靶的事件和 16 题中张三赢棋的事件．对于小概率事件，由于在一次试验中发生的可能性很小，因此，在一次试验中，实际上可把它看成是不可能发生的，我们称它为小概率事件实际不发生原理．但是，在大量试验中，小概率事件几乎必然发生．例如某事件发生的概率为 p，其中 $p<1$，则在 n 次独立重复试验中，该事件发生的概率为 $P=1-(1-p)^n$，显然当 $n\to+\infty$时，$P\to 1$.

17. 不可能．因为他输赢的概率为 0.5，由频率与概率的关系，大量赌博后输赢的次数大致差不多，因此不可能赢到 2 万元.