高等教育“十二五”规划教材

公共课教材系列

马元生 主 编

杨时中 主 审

概率统计简明教程（第二版）

GAILV TONGJI JIANMING JIAOCHENG

科学出版社

高等教育"十二五"规划教材

公共课教材系列

概率统计简明教程

（第二版）

马元生　主　编

曹黎霞　高　梅　副主编

杨时中　主　审

科学出版社

北　京

内 容 简 介

本书以教材应为学生服务、应能起到教的作用为宗旨，本着在教材中落实对学生素质能力培养的意愿，在教学方法上采取了向学生演示解惑过程的方式，即以提出问题、研究解决问题为主线，经过分析、归纳、提炼、概括形成所要讲述的概念或定理，让学生看到知识形成的过程，从中学习解惑的方法，提高自身的解惑能力. 本书注重对概念理论的理解和结论的应用.

本书介绍了概率论的基本知识和常用数理统计方法与原理，可作为普通高校、成人教育、自学考试理工类及经管类专业的教学用书.

图书在版编目(CIP)数据

概率统计简明教程/马元生主编. —2版. —北京:科学出版社,2011
(高等教育"十二五"规划教材・公共课教材系列)
ISBN 978-7-03-031079-8

Ⅰ.①概… Ⅱ.①马… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ.①O21

中国版本图书馆CIP数据核字(2011)第088987号

责任编辑:王 彦 / 责任校对:柏连海
责任印制:吕春珉 / 封面设计:东方人华平面设计部

科学出版社出版
北京东黄城根北街16号
邮政编码: 100717
http://www.sciencep.com
源海印刷有限责任公司 印刷
科学出版社发行 各地新华书店经销
*
2007年9月第 一 版 开本: B5(720×1000)
2011年6月第 二 版 印张: 14 1/4
2011年6月第六次印刷 字数: 271 000
印数:12 501—15 500
定价: 25.00 元
(如有印装质量问题,我社负责调换〈路通〉)
销售部电话 010-62136075 编辑部电话 010-62138978-8208

第二版前言

2011年是笔者从事教学工作的第五十个年头，半个世纪以来与知识和学生打交道是很愉快的，但同时又感到“教”海深邃，当好一名教师是多么的不易. 教学就是要让学生学到知识. 一个从不知到知的过程中的规律性和具体的教学运作是非常精细、缜密，非常科学的. 谨就教材而言，教材是什么？教材要在这个过程中起到什么作用？笔者工作了45年方有所悟. 教材是教学方法与知识内容巧妙的地有机结合，教材本身就应当具有教的功能，发挥教的作用. 笔者的两本书《概率统计简明教程》和《线性代数简明教程》，便是本着上述的认识对知识内容作了许多处理加工写成. 一方面是为了学生学习，让中等水平的学生就能够读懂，但对概念、理论又要阐述到位，不失水准. 另一方面是为了让教师用起来方便顺手，无需再做大幅度的加工处理. 书中对许多概念、定理的引入和对内容的处理，体现了启发式教学方法和反映了解决问题的思想过程. 本书于2007年面世后，承蒙广大同仁的支持厚爱和学子们的欢迎，使笔者深受鼓励，特向同仁们深切致谢！第二版更正了原书中的许多错误，增加了前四章的习题，对概率分布和数学期望的阐述更加贴切. 为节省课时方便教师安排教学进度，对内容稍作调整.

本书由马元生教授（西安工业大学、西安思源学院）担任主编，曹黎霞、高梅担任副主编. 本书编写分工如下：柴伟文（第1、3、4章）、曹黎霞（第2、5、6章）、高梅（第7至第9章）.

由于编者水平所限，错误仍在所难免，深望同仁们批评指正.

编　者

2011年2月

第一版前言

教学是大家熟悉的工作.教,就是要讲出道理来,讲不出道理,教就不到位,所以教师的工作是边做学问边教知识.即使是教知识,也不能是单纯的传授知识,而应注重培养学生的素质和能力.也就是说,教师的工作不只是传道、授业和解学生之惑,而更重要的是通过向学生讲授知识,教给学生解惑的能力,传授给学生解惑的本领,让学生能够在日后的工作中会发现问题和研究、解决问题.那么教师怎样才能教学生解惑呢?师傅带徒弟,只能由师傅先做给徒弟看,所以只能是由教师在学生面前演示解惑的过程.教师在所讲授的课程中,在难点处演示解惑的过程,各门课程的教师都努力这样去做,经过几年的熏陶和教育,会使学生的思维方法、思维模式、思维能力有所提高,学生自然会掌握一些解惑的方法和本领,素质和能力必然会有所提高.

数学课教师怎样向学生演示解惑呢?这个问题涉及讲课模式."定义、定理、证明、举例"是数学课的传统讲课模式.如果讲定义,教师是先写出定义,然后解释定义;讲定理,教师只是解释定理的条件和结论,然后即开始证明;讲证明不先分析内在联系、因果关系理出思路,而是讲我们第一如何……第二如何……这样讲课就没有向学生演示解惑的过程.上述的讲课模式并不是不对,关键是教师怎样讲.笔者认为讲概念和定理都要讲背景,要以提出问题,研究、解决问题的方式展开内容.要讲从研究什么问题出发……经过分析研究使问题得到解决.通过对这一过程的分析、归纳、提炼、概括,形成新的概念或定理的雏形,所要介绍的定义、定理自然地顺势而出.这样讲授可以向学生再现知识的发现和形成的过程,让学生看到知识形成的轨迹,同时会感到所引入的概念和定理是自然的、合理的,是因为"需要"才引入的,概念和定理成了解惑的结果.能够这样讲课,可以说就有了演示解惑的味道.

以上表述了笔者对教学工作的认知,本教材就是在上述认知的基础上,笔者在教学过程中的努力与实践的结果.如果读者对以上的表述给予认同,相信本书对您会是有益的.本书注重对概念、理论的理解和结论

的应用,而不拘泥于理论的严谨求证.对概率论的内容作了一些处理,笔者认为一维随机变量是概率论的重点,讲透彻一维,多维便容易学习.所以将一维随机变量的分布及期望、方差作为一章.将二维随机变量的分布及协方差、相关系数作为一章.在数理统计部分着重讲清各种统计方法的原理和操作步骤,便于学生掌握运用.

本书介绍了普通高等院校理工科概率统计课程的全部内容.由于概率统计是一门比较难学的数学课,所以笔者注重对概念、定理的引入和对概念、定理的理解并将习题归类,写得深入浅出、通俗易懂,对高职高专学生也是很适宜的.每章的小结阐明了如何去理解看待本章的内容,对同学们是有益的.

鉴于编者水平有限,书中错误在所难免,望各位同仁批评指正.

最后要感谢科学出版社的万国清、王彦、游浩星三位老师,他们理解作者从事多年教学工作的认识和体会,积极鼓励作者将此融于教学内容之中写成教材,以对学生有所帮助,并且在酷暑中认真地进行了编辑工作,在此,作者谨向三位老师表示衷心的感谢!

目　录

预 备 知 识

一、乘法原理

例 0.1 设由甲地到乙地有 2 条路可走，由乙地到丙地有 3 条路可走. 由甲地到丙地必须经过乙地. 问由甲地到丙地共有多少种路线可走?

解 设由甲地到乙地的两条路分别为I_1，I_2. 乙地到丙地的 3 条路分别为II_1，II_2，II_3. 那么由甲地到达丙地需要两段路的搭配，共有

$$(\mathrm{I}_1,\mathrm{II}_1),\quad (\mathrm{I}_1,\mathrm{II}_2),\quad (\mathrm{I}_1,\mathrm{II}_3)$$
$$(\mathrm{I}_2,\mathrm{II}_1),\quad (\mathrm{I}_2,\mathrm{II}_2),\quad (\mathrm{I}_2,\mathrm{II}_3)$$

6 种不同的路线.

由甲地经过乙地而到达丙地分两个步骤，先从甲到乙，完成方式有两种：I_1，I_2；再由乙地到丙地，完成方式有 3 种：II_1，II_2，II_3. 完成甲地经乙地到丙地总的路线数目等于这两段完成方式数目的乘积. 一般有：

完成一项工作需 n 个步骤，而完成每个步骤的方式分别为 $m_1,m_2,\cdots,m_n$ 种，则完成这项工作的方式共有 $m_1\cdot m_2\cdot\cdots\cdot m_n$ 种.

称此结论为乘法原理.

例 0.2 用数码 1,2,3,…,8,9 写成三位数(数码不重复)，试问能写出多少个不同的三位数?

解 写出一个三位数需 3 个步骤，先写百位数，再写十位数，最后写出个位数. 而完成每个步骤的方式易知分别为 9,8,7 种. 所以，按乘法原理共能写出 $9\cdot 8\cdot 7=504$个不同的三位数.

二、排列

1. 全排列

按乘法原理的思考方法，易知 n 个不同元素的全排列数为 $n!$ 个.

2. 选排列

易知从 n 个不同的元素中取 m 个元素的选排列数为

$$n\cdot(n-1)\cdot(n-2)\cdot\cdots\cdot(n-m+1),$$

记为

$$A_n^m=n\cdot(n-1)\cdot\cdots\cdot(n-m+1).$$

3. 可重复排列

例 0.3　由数码 1,2,3,…,8,9 写成的三位数(允许重复)有多少个?

解　写出一个三位数仍需 3 个步骤,而由于可以重复,所以每个步骤的完成方式都为 9,所以可写出 $9^3=729$ 个三位数.

一般有,从 n 个不同元素中取出 m 个元素的可重复排列数为 n^m. 记为

$$r_n^m = n^m.$$

在此要强调,排列的要素有两个,一是这个排列由哪些元素构成,二是这些元素排列的顺序.

例如,对于以下 3 个排列,

$$(1,2,3,4) \quad ①$$

$$(1,2,3,5) \quad ②$$

$$(1,2,4,3) \quad ③$$

排列①与排列②不同,是由于其所含的元素不完全相同;排列①与排列③不同,在于元素的排序不同.

三、组合

例 0.4　从数码 1,2,3,…,8,9 中取 4 个数码,试问有多少种不同的取法(或结果).

注意　这里只是从 9 个不同的元素中取出 4 个元素,而不需再排列,这 4 个元素就是从 9 个元素构成的集合中取出 4 个元素的一个组合,所以“组合”就是原来集合的一个子集. 因此组合的要素只有一个,即构成此组合的元素.

现在我们来考虑上面例题的解,由于组合对于我们是新问题,而排列已经是熟悉的解决了的问题,所以,我们可以考虑借助排列来解决组合数的问题.

我们将从 9 个数码 1,2,3,…,8,9 中取 4 个数码的选排列的全体进行分类,把所含 4 个元素皆相同的排列归为一类. 那么每一类中排列的个数应是 4 个元素的全排列数 4! 个. 而不同类的排列其所含的元素不尽相同,所以进行分类的类数即是从 9 个元素中取 4 个元素的组合的数目,记为 C_9^4. 所以有

$$C_9^4 \cdot 4! = A_9^4,$$

$$C_9^4 = \frac{A_9^4}{4!} = \frac{9 \cdot 8 \cdot 7 \cdot 6}{4!} = 126\ .$$

一般情形有:

从 n 个不同元素中取 m 个元素的组合数目

$$C_n^m = \frac{n(n-1)\cdots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}.$$

例 0.5 某一新铁路干线即将完工，现准备运营印制车票，铁路线上的站名共 10 个，试问应印制多少种车票？又有多少种票价（两站之间往返票价相同）？

解 印制车票种类数

$$A_{10}^2 = 10 \cdot 9 = 90 ,$$

票价种类数为

$$C_{10}^2 = \frac{10 \cdot 9}{2!} = 45 .$$

注意 关于组合数有两个常用的等式

$$C_n^m = C_n^{n-m},$$

$$C_n^m = C_{n-1}^{m-1} + C_{n-1}^m.$$

希望读者勿需推证，而要悟出其中的道理，从而两个等式成立是很自然的事.

第 1 章　随机事件与概率

§1.1　随机事件与概率

在日常生活和工程实践中会遇到很多事物或现象，我们称这些事物、现象为事件. 就事件发生或不发生的情况而言是有差别的. 例如，温度降至 0℃以下，水一定会结冰；温度升至 100℃，水一定会沸腾；重物失去支撑一定会下落. 所述 3 个例子中的事件发生的情况有一个共同之处，即当所述的条件具备，事件一定发生. 或者说这种试验进行多少次，相应的事件就发生多少次. 我们称这类事件为必然事件. 另一类事件发生的情况恰好与之相反，例如温度未降至 0℃，水结冰是不可能发生的. 温度未升至 100℃，水沸腾是不可能发生的. 重物有支撑，重物下落是不可能发生的. 这类事件发生情况的一个共同特点是，只要进行试验，该事件一定不发生. 我们称这类事件为不可能事件. 必然事件和不可能事件就其发生的情况而言都是确定性事件.

另外还有一类事件，其发生或者不发生都是不确定的. 例如，购买一张中奖的彩票；从一副扑克牌中任取一张为黑桃 A；明天下雨. 这些事件就其发生的情况也有共同点，即如果进行试验，事件有可能发生，但也可能不发生. 即这类事件的发生与否带有机会性. 机会好，购买的彩票中奖；机会不好，购得的彩票不中奖. 我们称这类事件为随机事件. 如果进行试验，这种事件可能发生，也可能不发生. 以后用字母 A,B,C 表示随机事件.

任何随机事件都伴随着有相应的试验，我们称这种试验为随机试验，记为 E. 它具有如下 3 个特点：

（1）在试验之前知道（或明确）如果进行试验将会出现的各种可能的结果.

（2）试验之前不能确定试验之后会出现哪一个结果.

（3）能够具备相同的条件，使试验可以重复进行多次.

以上列举的购买彩票、抽牌都是随机试验. 某人向一目标进行射击也是随机试验，因为射击前就知道，如果进行射击可能出现的结果一个是击中目标，另一个是未击中目标. 但射击前不能确定发生哪一个结果. 可以重复多次进行射击. 所以某人向目标进行射击是随机试验.

根据以上的叙述，我们应该有两个认知：

（1）试验结果具有排他性，即试验结果之间清晰可辨，不会发生亦此亦彼的情况.

(2) 每进行一次试验，只有一个试验结果出现，不会有不同的试验结果同时出现.

我们将随机试验 E 所有可能的结果 $e_1,e_2,\cdots$ 组成的集合称为 E 的样本空间，记作 $\Omega,\Omega=\{e_1,e_2,\cdots\}$. 将 $e_1,e_2,\cdots$ 各自组成的单元集称为 E 的基本事件. 为了避免符号过于复杂，仍用 $e_1,e_2,\cdots$ 表示. 即符号 $e_i(i=1,2,\cdots)$ 既可表示 E 的结果，又可表示 E 的基本事件.

例 1.1.1　从 13 张黑桃牌中任取一张. 这个随机试验的试验结果有 13 个：黑桃 A，黑桃 2，黑桃 3，…，黑桃 K. 可以用 $e_i(i=1,2,\cdots,13)$ 分别表示相应的试验结果. 那么这个随机试验的样本空间

$$\Omega=\{e_1,e_2,\cdots,e_{13}\}\ .$$

有了基本事件和样本空间的概念，试验中的随机事件便可以得到更具体的描述.

例 1.1.2　从 13 张黑桃牌中任取一张，令 A、B、C 分别表示如下事件：

A：抽到的牌为偶数牌.

B：抽到的牌为奇数牌.

C：抽到的牌上号码不小于 10，且为偶数牌.

分析事件 A、B、C 各自内涵的构成.

对于事件 A，只要抽到的牌上的号码是偶数，A 就发生. 所以样本空间中对 A 的发生有利的场合是 $e_2,e_4,e_6,e_8,e_{10},e_{12}$ 等试验结果. 也可以说事件 A 包含了这些试验结果. 对此我们表示为

$$A=\{e_2,e_4,e_6,e_8,e_{10},e_{12}\}.$$

对上式应理解为：右端集合中的任一试验结果发生，则事件 A 发生. 反之，A 发生，则右端集合中必有某一个试验结果发生. 这样，我们就把试验中的随机事件 A，看成了由试验的某些结果构成的集合. 即将事件 A(的构成)看成是试验样本空间 Ω 的一个子集合. 据此，有

$$B=\{e_1,e_3,e_5,e_7,e_9,e_{11},e_{13}\}\ ,$$

$$C=\{e_{10},e_{12}\}\ .$$

以及试验中的必然事件是 Ω，不可能事件是空集 $\varnothing$.

因此我们对随机事件的认识应当是：

(1) 事件的发生或不发生都带有随机性.

(2) 从构成上来看，随机事件是样本空间的子集合.

(3) 事件发生的频率具有稳定性.

上述的第 3 条是基于人们的实践经验. 如在本章第 2 节中的重复投币试验，统计正面向上的频率，确有随着试验次数的增大而频率逐渐稳定在某个数附近的事实. 这一条也称作随机事件具有统计规律性. 可以这样说，目前人们能够处理的随机现象，都是具有统计规律性的随机现象. 随机性太强，例如从楼顶上向地面抛气球，统计气球的落地位置，可能抛一万次会有一万个不同的落点. 对于这样的随机

现象目前人们还不能处理.

在例 1.1.2 中的事件 A、B、C 之间会有某些关系. 例如，事件 A 发生，则事件 B 就不发生. 反之，事件 B 发生，事件 A 就不发生. 事件 C 发生，则事件 A 发生. 在一个随机试验中会有很多随机事件，弄清它们之间的关系是必要的. 下面就来讨论事件关系与运算.

为了直观，我们画出了图 1.1～图 1.3，图中的大方框 Ω 表示跳伞运动员可能落入的场地. 跳伞员每次跳伞都可能落入场地 Ω 内的任何一个位置，但不会落在 Ω 外，所以落入 Ω 内是必然事件. 在场地 Ω 内，又画出了 A 圈、B 圈，我们约定跳伞员落入 A 圈称事件 A 发生，落入 B 圈称事件 B 发生.

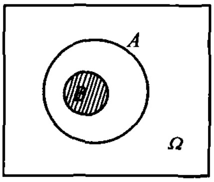

图 1.1

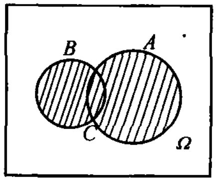

图 1.2

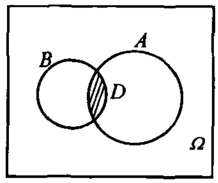

图 1.3

(1) 子事件.

如图 1.1 所示，设跳运动员落入 B 圈，事件 B 发生. 但这时运动员自然落入 A 圈，导致事件 A 发生. 事件 A,B 间有关系：

事件 B 发生则导致事件 A 发生，称 B 为 A 的子事件，记为 $B \subset A$.

例 1.1.2 中的事件 C 与 A 的关系，即 C 是 A 的子事件.

(2) 相等事件.

如果 A 圈与 B 圈重合，则事件 A,B 互为子事件.

若事件 A,B 互为子事件，称 A,B 为相等事件，记为 $A=B$.

(3) 和事件.

在事件 A,B 的基础上构成另一新事件 C，C 与 A,B 的关系是：A 发生，或者 B 发生，则 C 发生. 或者，若 C 发生，则 A,B 至少其一发生. 称 C 为事件 A,B 的和事件，记为 $C=A+B$.

如果 A,B 的和事件 $A+B$ 未发生，则 A,B 皆未发生.

如图 1.2 中 A 圈，B 圈所围场地的外轮廓曲线 C 所围之场地，跳伞运动员落入其中，即 A,B 的和事件 $C=A+B$ 发生.

(4) 积事件.

由事件 A,B 构成另一新事件 D，当 A 发生而且 B 发生，则事件 D 发生，称 D 为 A,B 的积事件. 记为 $D=AB$.

如果事件 A,B 至少其一未发生，则积事件 AB 不发生. 如果积事件 AB 发生，

则 A 和 B 都发生.

如图 1.3 中 A,B 相交的部分,跳伞运动员若落入其内,则 AB 发生.

(5) 差事件.

由事件 A,B 又构成一个新事件 F,表示 A 发生但 B 不发生. 事件 F 称作 A 与 B 的差事件,记作 $F=A-B$.

(6) 互不相容(互斥)事件.

事件 A,B,若其中一个发生则另一事件不发生,称 A,B 为互不相容(或互斥)事件,表示为 $AB=\varnothing$.

例 1.1.2 中的事件 A,B 是互不相容的. 事件 B,C 也是互不相容的.

注意　例 1.1.2 中的事件 A,B,

$$A=\{e_2,e_4,e_6,e_8,e_{10},e_{12}\},$$
$$B=\{e_1,e_3,e_5,e_7,e_9,e_{11},e_{13}\},$$

由于 A,B 互不相容,所以 A,B 不含相同的试验结果,分析事件 $A+B$ 的构成则有

$$A+B=\{e_1,e_2,e_3,e_4,\cdots,e_{12},e_{13}\}=\Omega.$$

上式表明 A,B 的和事件是必然事件,即只要进行试验,$A+B$ 就发生. 但由于 A,B 互不相容,所以只能出现 A,B 中的一个事件. 即只要进行试验,事件 A,B 中有且仅有一个发生,我们将 A,B 之间的这种关系称为对立事件.

(7) 对立事件.

事件 A,B 互不相容,且其和事件为必然事件,称 A,B 为对立事件. 亦称 B 为 A 的逆事件,记作 $B=\overline{A}$.

所以

$$A\overline{A}=\varnothing,$$
$$A+\overline{A}=\Omega.$$

事件之间的运算满足一些熟知的运算规律:

(1) 加法及乘法的交换律.

$$A+B=B+A,\ AB=BA.$$

(2) 加法及乘法的结合律.

$$A+(B+C)=(A+B)+C,\ A(BC)=(AB)C.$$

(3) 分配律.

$$A(B+C)=AB+AC.$$

(4) 对偶律(摩根定律).

$$\overline{A+B}=\overline{A}\,\overline{B},$$
$$\overline{AB}=\overline{A}+\overline{B}.$$

注意　上述这些等式的两端都是事件,所以每个等式的两端为相等事件,如果要进行证明,则要按相等事件概念的要求去做,即证明等号两端的事件互为子事件

方可. 下面仅以证明对偶律的第一个为例,其他,读者可自行练习.

试证明
$$\overline{A+B}=\overline{A}\,\overline{B}.$$

证 设左端$\overline{A+B}$发生,则 $A+B$ 不发生,故其两项 A 和 B 皆不发生. 但由此,$\overline{A}$ 且 $\overline{B}$ 皆发生,所以 $\overline{A}\cdot\overline{B}$ 发生. 故有$\overline{A+B}\subset\overline{A}\,\overline{B}$.

反之,设 $\overline{A}\,\overline{B}$ 发生,则 $\overline{A}$ 和 $\overline{B}$ 皆发生. 因此 A 和 B 皆不发生,所以 $A+B$ 不发生,而其逆事件$\overline{A+B}$发生,故有 $\overline{A}\,\overline{B}\subset\overline{A+B}$.

所以
$$\overline{A+B}=\overline{A}\,\overline{B}.$$

作为本节的结束,简单介绍一下随机事件的概率的概念.

例 1.1.2 中的事件 A,B,C 在抽牌试验中都有发生的可能性. 再进一步思考一下,它们发生的可能性的大小是有差别的. 事件 B 发生的可能性最大,事件 A 发生的可能性次之,事件 C 发生的可能性最小. 怎样来表现、反映事件发生的可能性的大小不同,自然想到是用数来反映. 即

在试验中事件发生的可能性是客观存在的,且有差异. 用数来反映事件发生的可能性,称此数为事件的概率.

大家熟悉长度、面积、体积等概念. 这都是属于度量性的概念,如果没有这些概念,跑百米和跑马拉松就无法区别了. 这 3 个概念分别是对具有线、面、体几何特征的物体的度量性概念,而且其度量结果用数表示. 同样,概率也属度量性概念,不过其度量的对象是事件发生的可能性. 这种可能性看不见摸不着,但可以感觉,可以领悟到它的存在. 对事件发生的可能性进行度量和计算,较之对几何物体的某种度量和计算显得更抽象更困难一些,需要我们付出更多的努力.

§1.2 频率与概率,古典概型中概率的计算

事先规定好一枚硬币的一面为正面,另一面为反面,在相同的高度掷向桌面,记录它向上的一面是正面,或者是反面. 历史上许多统计学家作过多次这种掷硬币的试验,表 1.1 汇集了蒲丰(Buuffon),皮尔逊(Pearson)等 4 位统计学家重复投币的统计情况.

表 1.1

试验者	投掷次数 n	出现正面向上次数 k	出现正面向上的频率$\frac{k}{n}$
摩　根	2 048	1 061	0.518 0
蒲　丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5
维　尼	30 000	14 994	0.499 6