应用统计学系列教材 Texts in Applied Statistics

抽样技术及其应用

Sampling Techniques and Practices

杜子芳 编著

Du Zifang

清华大学出版社 Springer

应用统计学系列教材 Texts in Applied Statistics

抽样技术及其应用

Sampling Techniques and Practices

杜子芳 编著

Du Zifang

清华大学出版社 Springer

北京

内容简介

全书共分9章，第1章介绍了学习抽样理论需要的预备知识；第2章介绍了抽样理论的若干基本概念；第3章到第7章分别对常规的几种抽样方法——简单随机抽样、分层随机抽样、二阶及多阶抽样、整群抽样和系统抽样，围绕其基本概念、特点、适用场合、具体抽样步骤和相应的简单估计、比率估计和回归估计三种估计量的性质等几个方面，进行了深入详尽的介绍，并且特别对每一种抽样方法的三种估计量的性质融会在一起进行讨论、比较；第8章介绍了实际中极为常见的非概率抽样的各种方式，重点阐释了这些方式与概率抽样的几种抽样方法间的对应关系；第9章是关于与抽样有关但又不属于抽样的一些问题的讨论，帮助实际工作者开阔眼界，避免失误. 出于同样的目的，在本书的重要章节还至少给出了一个相应的案例.

本书可作为统计专业的本科生、非统计专业的研究生教材，也可供相关专业的教师及实际从事抽样工作的人员参考.

图书在版编目(CIP)数据

抽样技术及其应用/杜子芳编著. —北京：清华大学出版社，2005.8
(应用统计学系列教材)
ISBN 7-302-11039-5

Ⅰ.抽…　Ⅱ.杜…　Ⅲ.抽样调查－高等学校－教材　Ⅳ.C811

中国版本图书馆CIP数据核字(2005)第050112号

出 版 者：清华大学出版社　　**地　　址**：北京清华大学学研大厦
http://www.tup.com.cn　　**邮　　编**：100084
社 总 机：010-62770175　　**客户服务**：010-62776969

责任编辑：王海燕
封面设计：常雪影
印 装 者：清华大学印刷厂
发 行 者：新华书店总店北京发行所
开　　本：170×230　**印张**：32.75　**字数**：586千字
版　　次：2005年8月第1版　2005年8月第1次印刷
书　　号：ISBN 7-302-11039-5/O·463
印　　数：1～3000
定　　价：46.00元

应用统计学系列教材

Texts in Applied Statistics

编审委员会

序

随着社会经济的飞速发展，统计学课程设置的不断调整，统计学教材已经有了很大的变化。为了适应这些变化，我们从2000年开始编写面向21世纪统计学系列教材，经过近4年的实践，该系列教材取得了较好的效果，基本实现了预定的目标。然而目前学科的发展和社会的进步速度相当快，其中的一些教材已经需要进一步修订，也有部分内容成熟、适合教学需要的教材没有列入编写计划。

为满足应用统计科学和我国高等教育迅速发展的需求，清华大学出版社和施普林格出版社(Springer-Verlag)合作，倡议出版这一套“应用统计学系列教材”，作为对现有统计学教材的全面补充和修订。这套教材具有以下特点：

1. 此套丛书属于开放式的，一旦有好的选题，即可列入出版计划。

2. 在教材选择上，拓宽了范围。有些教材主要面向经济类统计学专业，包括金融统计、风险管理与精算方面的教材。部分教材面向人文社科专业，而另外一些教材则面向自然科学领域，包括生物统计、医学统计、公共卫生统计等。

3. 本套教材的编写者都是活跃在教学、科研第一线的教师，他们能够积极地、广泛地吸收国内外最新的优秀成果。能够在教学中反复对教材进行补充修订和完善。

4. 强调与计算机应用的结合，在教材编写中，注重计算机软件的应用，特别是可编程软件的应用。对于那些仅限于应用方法的教材，充分考虑读者的需求，尽量介绍简单易学的“傻瓜”

软件。

5. 本套教材包括部分优秀国外教材译著，对于目前急需，而国内尚属空白的教材，选择部分国外具有广泛影响的教材，进行翻译出版。

我们希望这套系列教材的出版能够对我国应用统计科学的教育和我国统计事业的健康发展起到积极作用。感谢参与教材编写的中国人民大学统计学院和兄弟院校的教师以及进行审阅的同行专家。让我们大家共同努力，创造我国应用统计学科新的辉煌。

易丹辉

2004 年 1 月

前　言

由于决策科学化程度的提高，调查这种系统地搜集信息的活动已经成为各种研究的基础，“没有调查就没有发言权”的法则现今在任何领域、任何层面均被普遍认同和遵守。而在任何较大规模的调查中，抽样技术的选择与应用皆是不可避免的。不幸的是，采用抽样技术的做法的普及反而凸显了相关知识普及的不足，除了某些官方机构和学术机构组织的抽样调查比较规范以外，即使那些在大众媒体上亮相频繁、抢尽风头的专业调查公司所公布的抽样方案，在科学意义上也很少称得上规范合格。

笔者在中国人民大学讲授抽样理论十几年，期间也断断续续主持或参与了不少政府或企业的调查设计工作。长期的教学与实践使我感受到，虽然前辈同行迄今做出了巨大努力：国际上一些抽样方面的名著悉数得到翻译，国内也出版了一些专著和教科书，但所有这些相对于人们对抽样技术理论和方法的了解的迫切需要程度来讲，仍相去甚远。坦率地说，国内现存的有关抽样技术的书籍，要么过于专业，内容艰涩难懂；要么过于业余，内容零碎不成系统，其中少数出版物更满是将“总体”混同“人口”之类的低级错误，使人愕然。

鉴于此，本书主要面向那些在大学里讲授抽样技术课程的教师、统计专业的本科生、非统计专业的研究生和相关部门的中高级人士。内容力求详尽而不失精练，结构力求系统而不失紧凑。理论证明步骤明晰，浅显易懂，对数理统计知识要求不多，以便教师备课和读者自学，例题、习题选题广泛，兼顾明理与实用的双重需要。

全书共分9章，第1章介绍了学习抽样理论需要的预备知识，包括排列组合、中心极限定理、参数的区间估计和调查概论；第2章介绍了抽样理论的若干基本概念；第3章到第7章分别对常规的几种抽样方法——简单随机抽样、分层随机抽样、二阶及多阶抽样、整群抽样和系统抽样，围绕其基本概念、特点、适用场合、具体抽样步骤和相应的简单估计、比率估计和回归估计三种估计量的性质等几个方面，进行了深入详尽的介绍。特别地，对每一抽样方法的三种估计量的性质融会在一起进行讨论、比较，构成了本书的特色之一。本书的另一个特色是对不等概率抽样的处理，对其不做独立的讨论，但强调在次级或基本单元的层次上仍属于等概率抽样的特点，这样有助于读者更好地理解所谓不等概率抽样的实质，避免无谓的混淆。第8章介绍了实际中极为常见的非概率抽样的各种方式，重点阐释了这些方式与概率抽样的几种抽样方法间的对应关系。这个问题是极其重要的，可惜国内外绝大多数抽样文献都未给予必要的重视。最后一章是关于与抽样有关但又不属于抽样的一些问题的讨论，帮助实际工作者开阔眼界，避免失误。出于同样的目的，在本书的重要章节还至少给出了一个相应的案例。

在本书写作过程中，我的一些研究生朱华鹏、姜莉莉、江智杰、徐瑶、徐晓菊、牛奇和王欣等同学给予了许多帮助，他们不仅参与了部分章节的撰写、习题案例的寻找，还帮助调整了全部的格式。特别是朱华鹏和姜莉莉，对本书的结构提出了很多宝贵建议，又反复阅读和校对了全部书稿，保证了稿件的按时完成。在此谨表示诚挚的谢意。最后，作者还要特别感谢我的同事张波教授，他对本书的顺利完成助益良多，付出了很多心血与努力。

由于作者水平有限，写作亦较仓促，虽然诚惶诚恐，殚精竭虑，但不足之处仍然在所难免，甚至可能不少，敬请有关专家、同行和广大读者赐教、斧正。

杜子芳

2005年2月

目 录

第1章

预备知识

本章将简要地介绍排列组合、概率统计、调查等方面的有关知识.这些知识属于抽样技术的基础知识或预备知识,在以后各章的叙述中会用到,但又不属于本书中任何其他章节的介绍范围.补充这些知识供读者在学习时参考,能够使读者更加顺利地学好以后各章节的内容.建议读者特别是那些没有系统学习过调查和数理统计的读者,对本章所介绍的内容能够认真阅读,因为这些内容不仅是进一步学习抽样技术的准备性知识,同时这些知识本身也是十分重要和有用的.

1.1 排列组合

1.1.1 基本原理

首先叙述两条基本原理,这两条原理在排列组合分析公式的推导中起着重要作用.

(1) 加法原理:假如完成一件事有 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,而无论通过哪种方式方法都可以完成这件事,则完成这件事总共有 $n_1+n_2+\cdots+n_m$ 种方法.

(2) 乘法原理:假如完成一件事有 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,而每一步骤都不可缺少也不可重复,则完成这件事共有 $n_1n_2\cdots n_m$ 种方法.