

现代人口分析技术

翟振武等著



现代人口分析技术

翟振武 路 磊 罗茂初 曲海波 编著

中国人民大学出版社

现代人口分析技术

翟振武 路 磊 编著
罗茂初 曲海波

●
中国人民大学出版社出版发行

(北京西郊海淀路39号)

中国人民大学出版社印刷厂印刷

(北京鼓楼西大石桥胡同61号)

新华书店经销

●
开本: 850×1168毫米32开 印张: 14.125

1989年12月第1版 1989年12月第1次印刷

字数: 346 000 册数: 1 - 3 000

●
ISBN 7-300-00610-8

C·43 定价: 5.60元

序

70年代以来,在我国人口工作实践不断深入的基础上,人口学也获得了历史上空前未有的发展。为了制定符合我国国情的有计划地控制人口增长的战略,人口学工作者研究和回答了实践中提出的大量理论问题,人口理论有了重大的突破和发展。与此同时,为了更深入地掌握我国的基本国情,进行了10亿人口普查、千分之一生育率调查、人口自然变动抽样调查、残疾人调查等。在此基础上,对调查资料进行了多方面研究分析,对人口发展进行了多种预测,展望未来可能出现的趋势。所有这些,使人口计算、人口普查、人口调查和人口统计分析也有了发展。正由于人口工作实践和理论的发展,80年代以来,大量的人口学教材、著作如雨后春笋般地相继问世。现在摆在读者面前的《现代人口分析技术》一书为人口学著作增添了新的篇章。它是广大人口学理论工作者和实际工作者早就期待出版的书籍。

1986年,我建议和鼓励本书作者们尽最大努力把这本著作奉献给人口学界、经济学界、社会学界和广大社会科学研究工作者和实际工作者。现在,这本书终于完成并出版,我感到非常高兴。

我为什么鼓励和建议他们完成这本著作呢?这是基于以下几方面的考虑:

(一)人口学科建设的需要。我认为,人口学这门学科包括三大组成部分:理论,方法和各应用分支学科。人口理论著作无论在数量上和质量上已有一定基础。在人口学方法方面:人口统

计学方法、人口调查、人口普查、人口预测等方面的教材或著作也出版了不少。但是比较全面、系统地介绍和论述人口分析方法，特别是70年代以后新发展出来的各种分析方法，如生育率模型、死亡率模型、间接估计方法等教材和著作却比较少见。因此，编写一本介绍现代人口分析技术的教材无疑是完善人口学学科建设的需要，它对培养社会科学人才必将起到应有的作用。当然，在人口学学科建设中还有许多空白要填补，任务极为艰巨，需要人口学工作者继续努力。

（二）人口实践工作的需要。近年来，为了了解和掌握我国各方面的基本情况，不同单位不同地区都进行了不少有关人口的调查。这些调查都取得了可贵的资料，为我们认识国情、制定具体政策提供了数据或依据。为了更加全面深入地分析这些数据和资料，编写一本有关如何用现代方法整理和分析调查所得数据的书籍是人口科学工作者一项迫切任务。因此，可以说这本书是适应我国人口实际工作需要的产物。

（三）进一步促进人口学发展的需要。人口学是一门综合性的社会科学。人口现象与社会、经济、文化、生态环境诸方面有着密不可分的联系。人口科学进一步丰富和发展除加强基础理论研究外，更有待于多方面的深入实际的定量研究和分析。我们既反对不重视定量研究分析，不从中国实际出发，从理论到理论的空洞议论，也反对只在数学模型上下功夫，不考虑它是否符合社会现象的运动规律、是否符合中国的实际而生搬硬套。我们只有以马克思主义为指导，把定量分析和定性分析结合起来，才能真正认识人口及其与社会、经济诸方面的内在联系，得出科学的认识，促进人口学的发展。《现代人口分析技术》一书的出版，将会在这方面起有益的作用。

本书作者翟振武、路磊和罗茂初是中国人民大学人口研究所研究人员，曲海波是博士研究生，他们都在国外接受过正规的人

口学培训，回国后，一直从事人口学的教学和研究工作。因此，他们既熟悉各种现代的人口分析技术，又了解中国人口的实际情况，这本《现代人口分析技术》就是他们辛勤工作的成果之一。我非常欣慰地看到新一代人口学者正在崛起，在教学和科学研究第一线上日益发挥越来越重要的作用，这是我国人口学兴旺发达的重要标志。我感到，我们年纪大一些的人口学工作者要给年轻的人口学者以积极支持，并创造必要的条件，充分发挥他们的才智和特长，推动人口学的繁荣发展，为培养人才和实现社会主义现代化服务。我也希望年轻的人口学者虚心向老同志学习，与人口学界老同志一道为共同发展我国人口科学贡献自己的力量。我想，这也是我国人口学界的共同愿望。

刘 铮

于1988年4月

前 言

人口分析技术是进行人口研究的工具。自70年代以来,随着人口问题越来越为世界瞩目,人口分析技术也获得了突飞猛进的发展。在数理人口学方面,稳定人口理论已被广义稳定人口理论所代替;在人口度量与分析上,各种新概念、新指标,如暂时平均预期寿命,寇尔生育指数等层出不穷;在人口模型方面,特拉赛尔模型生育率表、邦戈茨综合生育模型、罗吉特生命表系统等获得了日益广泛的应用;最令人刮目的是间接估计方法的发展,自从布拉斯等人在60年代末首次提出间接估计方法以来,短短的20年中,它已拓展为人口分析技术中一个崭新的独立领域,目前各种较为成熟的间接估计方法达几十种之多,在人口分析中发挥着越来越重要的作用。

中国有10亿人口,近几年经过人口普查、千分之一生育率调查、百分之一人口调查及各种各样人口抽样调查等,获得了极为丰富的人口资料与数据。因此,如何运用现代技术深入、全面地整理和分析这些资料与数据,便成为目前一项迫切任务。中国的人口学从70年代开始恢复以来,人口理论领域的发展很快,而人口分析技术领域的发展则相对薄弱一些,系统介绍和论述现代人口分析技术的著作也不多见,给我国人口数据的深入分析带来很多困难。《现代人口分析技术》正是在这样的背景下,为适应我国人口实践工作需要而编写的。

在编写过程中,我们针对中国人口的实际情况,全面、系统、深入地介绍和论述了可应用于中国人口实际数据的各种分析

技术，特别着重介绍和论述了70年代以来新发展出的人口分析指标、模型和方法。从内容上说，这本书基本囊括了人口学范围内各种现代分析技术。为了便于读者理解和运用这些技术，在编写方式上，我们不仅论述了各种分析技术的基本原理，假设条件和应用范围，而且尽可能详细地介绍了应用这些技术的具体步骤、实例和必要的系数参数表。因此，本书既可以作为社会科学工作者的参考书，作为人口分析技术培训课程的教材，也可以作为与人口有关的广大实际工作者的实用手册。

全书共分七章。考虑到数据评价是目前人口分析工作中的一个薄弱环节，第一章用大量的篇幅深入讨论了人口数据错误来源，数据质量评价方法以及如何调整数据等。第二章详细介绍了有关婚姻的度量与分析方法，并说明了编制结婚表的具体步骤。第三章和第五章分别介绍了生育与死亡的度量分析方法，包括了很多70年代以后才提出的新指标、新概念和新方法，对读者深入分析人口资料和数据极为有用。第四章和第六章则汇集了各种主要的生育率和死亡率的间接估计方法，各种生育模型、死亡模型及必要的系数参数表，它们可以帮助读者在缺乏资料的条件下，尽可能准确地估计和分析各种生育和死亡指标。第七章针对我国计划生育工作实际需要，专门介绍了各种有关计划生育的统计度量指标。

本书由作者集体讨论，分头执笔写成。第一章由罗茂初撰写；第二章、第七章由曲海波撰写；第三章、第四章由路磊撰写；第五章、第六章由翟振武撰写。初稿完成后，又经集体讨论修改，最后由翟振武对全书统一整理而成。

由于作者水平和时间的限制，书中一定有很多不妥甚至错误之处，我们诚恳地希望读者指正，以便再版时修改。

翟振武 路磊 罗茂初 曲海波

1988年4月

目 录

第一章 人口数据评价和调整	1
1.1 数据和数据质量	1
1.2 数据来源	4
1.3 错误来源	4
1.3.1 总误差模型	7
1.3.2 漏报和错报	10
1.4 数据评价	12
1.4.1 三种基本评价方法	13
1.4.2 年龄数据的质量及评价	17
1.4.3 出生、死亡数据评价	33
1.5 调整数据	42
第二章 婚姻度量与分析	55
2.1 婚姻的涵义及分类	55
2.2 婚姻状况资料的统计度量	57
2.2.1 用间接方法估算的初婚年龄中位数	57
2.2.2 用间接方法估算的平均初婚年龄	58
2.2.3 婚姻状况的图示	60
2.2.4 结婚年数	61
2.3 婚姻事件资料的统计度量	61
2.3.1 结婚诸率的度量	61
2.3.2 标准化方法的应用	65
2.3.3 结婚年龄的汇总与度量	67

2.4	结婚表的编制	68
2.4.1	结婚概率的计算	69
2.4.2	净结婚表的变量	70
2.4.3	净结婚表的编制	72
2.4.4	粗结婚表的变量和编制	78
2.5	离婚资料的统计度量	83
第三章	生育度量与分析	88
3.1	自然生育率	89
3.2	生育率的基本度量指标	92
3.3	标准化出生率和标准化生育率	110
3.3.1	直接标准化	111
3.3.2	间接标准化	113
3.4	生育次数递进比	116
3.5	寇尔生育指数	119
3.6	生育周期	124
3.7	生育率决定因素研究的框图	125
第四章	生育率间接估计方法	131
4.1	人口模型	131
4.1.1	寇尔-特拉赛尔生育模型	132
4.1.2	布拉斯生育率模型	139
4.1.3	邦戈茨综合生育率模型	145
4.2	根据儿童妇女人数估计生育率	157
4.2.1	儿童一妇女比方法	158
4.2.2	瑞利回归系数法	160
4.3	基于曾生子女数的生育率估计方法	161
4.3.1	P/F 比方法	164
4.3.2	曾生子女数增长法	172
4.4	存活倒推法	176

4.4.1	儿童存活倒推法	177
4.4.2	亲生子女法	181
4.5	根据稳定人口理论估计出生率	190
4.6	普莱斯顿出生率估计法	193
4.7	根据广义稳定人口公式估计普查间分年龄 生育率	200
第五章 死亡度量与分析		210
5.1	死亡率的标准化与分解	210
5.1.1	死亡度量的常用指标	210
5.1.2	死亡率的直接标准化	213
5.1.3	死亡率的间接标准化	218
5.1.4	死亡率比较指数	221
5.1.5	死亡率差别分解	225
5.2	死亡率变化对年龄结构的影响	229
5.3	婴儿死亡的度量与分析	237
5.3.1	婴儿死亡率	238
5.3.2	婴儿死亡分离系数	242
5.3.3	其他婴儿死亡水平的度量指标	245
5.3.4	婴儿死亡率变化对平均预期寿命的影响	248
5.4	死因统计及分析	251
5.4.1	死因资料的收集与分类	251
5.4.2	死因统计资料的分析方法	258
5.5	简略生命表	262
5.5.1	简略生命表各指标	263
5.5.2	生命表函数的几何解释	270
5.5.3	编制简略生命表的指数公式法	273
5.5.4	编制简略生命表的里德-麦瑞尔法	278
5.5.5	编制简略生命表的先行确定a值法	278
5.5.6	编制简略生命表的迭代法	280

5.5.7	两种调整生命表原始数据的方法	285
5.6	去某种死因生命表	288
5.7	平均预期寿命变化的度量与分析	295
5.7.1	暂时平均预期寿命	296
5.7.2	暂时平均预期寿命相对变化的测度指标	298
5.7.3	各年龄死亡率变化对平均预期寿命的影响	300
第六章	死亡率间接估计方法	308
6.1	模型生命表	308
6.1.1	1955年联合国模型生命表	309
6.1.2	寇尔-德曼区域模型生命表	310
6.1.3	莱德曼模型生命表	318
6.1.4	布拉斯罗吉特模型生命表	320
6.1.5	联合国发展中国家模型生命表	325
6.2	根据曾生子女尚存比例估计儿童死亡率	330
6.3	从孤儿比例数据中估计成人死亡率——孤儿法	344
6.3.1	布拉斯孤儿法	344
6.3.2	希尔-特拉赛尔回归系数法	354
6.4	从鳏寡比例数据中估计成人死亡率——鳏寡法	356
6.4.1	从分年龄鳏寡数据中估计成人死亡率	358
6.4.2	从按已婚年数分组的鳏寡数据中估计成人死亡率	363
6.5	从死亡年龄分布中估计成人死亡率	368
6.5.1	普莱斯顿-寇尔稳定人口法	369
6.5.2	增长平衡式法	381
6.5.3	贝内特-霍茹科广义稳定人口法	388
6.6	成人死亡概率与儿童死亡概率的衔接方法	391
6.7	从两次普查的分年龄人口数中估计死亡率	397
6.7.1	普查留存比法	398
6.7.2	普莱斯顿-贝内特广义稳定人口法	406

第七章 计划生育的统计度量	416
7.1 根据时点数据统计的诸指标	416
7.2 根据计划生育事件统计的诸指标	420
附 表	425
参考书目	433

第一章 人口数据评价和调整

现代人口分析非常注重数据的质量。在一定意义上讲,质量评价是人口分析的首要内容,只有对数据进行严格的检查,整理和说明后,分析工作才具有一个科学的坚实基础。在人口学体系内,数据评价是一个专门的领域,有其特定的理论、概念和方法,特别是第二次世界大战以后,随着世界范围内人口问题日趋引人关注,人口研究实践不断深化,这个领域内新概念、新技术不断涌现,其内容已变得非常丰富。

对发展中国家来说,人口数据的评价具有更重要的意义,因为这些国家大多数不具备完善、发达的数据收集系统,缺少一支高水平的调查统计队伍,许多这方面的尝试还刚刚起步,尚未取得足够的经验与成效。因此,人口研究工作者面对的常常是含有不同程度误差的数据,无法直接进入分析阶段,他们首先需要学习如何评价这些分析素材。

在介绍数据评价理论和方法之前,我们先介绍一些有关数据、数据来源和数据质量的概念,以求对人口分析的素材和评价标准有一全面、准确的了解。

1.1 数据和数据质量

数据有狭义和广义之分。在人口学领域内,狭义的数据仅仅指未经任何处理的人或事件的量,如人口数、出生数、节育数等等,一般是以绝对数的形式出现。广义的数据则不仅包括这些

量,而且包括由这些绝对数推导的人口学指标,如常用的出生率,死亡率,节育率等统计量,它们多是以百分比、比率、平均数等相对数的形式出现。

数据还可分为事实数据和意向数据两类。事实数据反映人的客观特征或所经历的有关事件,如年龄、文化程度、已婚妇女曾生育子女数。意向数据则反映人的价值取向及主观感受,如意愿生育子女数、对避孕措施的态度等。

数据(Data)和数字(Number)是有区别的。后者仅仅是前者的一种形式,有些数据以数字的形式出现,有些则不是,而是以文字形式出现,这在统计学上称为“列名变量”,其特点是列名变量的各范畴间不存在可相互运算或比较的关系。如“性别”变量,可分为“男性”和“女性”两个范畴,其间存在质的差别,不可以相互比较。

本章所介绍的数据评价原理和基本评价方法是针对广义数据的。凡人口学数据,不论是相对数、绝对数、事实数据、意向数据,还是数字、文字都在质量评价范围内,都需要对其质量作出严格的评定。但是对数据作以上的类型区分还是很重要的,因为它关系到总体评价技术的选择和质量标准的确立。

在社会科学范畴内,用来评判数据质量有多种标准,且容易相互混淆,有必要逐一加以讨论。

效度(Validity)。效度是指我们用来测度某一种现象或观念的标识是否得当,是否真确地反映了想研究的客观现实。因此,效度是衡量数据质量时首先要考虑的标准。效度不高,数字再准确也不能认为是质量好,因为它不符合研究的意图和目的。

信度(Reliability)。一旦测度确定后,数据质量的高低可以看其信度如何。信度又称“可靠性”,指“在同样条件下,重复测度某现象或观念,其结果的一致性”,从这个定义可以看出,信度是可以通过重复测度,从而从结果的一致性上加以判断

的，因此，它是就检测可能性而言的一个标准。从统计学的角度而言，信度反映的是数据的稳定性，即排除随机误差干扰的质量。

准确度 (Accuracy)。这是一个比信度更完整的标准，也是本章介绍的评价数据质量的准绳。所谓准确度，是指调查值与真实值之间的差距，差距越小，准确度越高。从理论上讲，数据准确，其信度一定高，但信度高却并不一定准确，因为信度是观察值自身比较的结果，而准确度反映观察值是否与真实值吻合。前者只能反映观察值中随机扰动的成份，后者不单包括随机扰动，而且包括非随机的偏差成份。

当我们使用这三个标准时，应注意到其概念的联系和区别。本章虽然主要介绍数据准确度的检验，但信度检验也是数据评价中一个很重要的部分，对信度的测定已有专门技术，如一致性试验 (Consistency Test)，读者可参考书后附录的参考文献。表1—1是印度尼西亚生育率调查数据的信度测试结果，表中右部为各类型数据信度的排序，比较起来，事实性的生育率数据信

表 1—1

数据类型		排 序
事实数据		
生育率数据	引导性问题	1
	数据型问题	2
背景数据	婚姻状况	3
	住址和职业	4
年龄和时间	年	5
	月	6
	日	7
意向和意见数据		8

来源：An Assesment of the Reliabilily of the Indonesia Fertily Survey, WFS Scienlific Report No. 3.

度最高，背景数据信度次之，年龄虽也属事实性数据，但呈现较低信度，最不稳定的是意向性数据。这种测试有助于我们了解不同类型数据的可靠程度，帮助我们改进调查问题的设计。

1.2 数据来源

人口数据可以从两种来源获得。

直接来源。指以专门收集有关人口数据为目的建立的来源，其主要手段包括人口普查、人口调查（抽样调查、典型调查）和经常性人口统计（又称生命统计或民事统计）。

间接来源。不是以专门提供人口数据为目的，但包括有关人口信息的来源。如国民经济、社会各部门的统计报表，调查报告，劳动人事部门劳动力统计资料，卫生部门的生育手术记录，医药销售部门关于避孕药具的销售、分配统计，民政部门对老年和残疾人保障金、补助金的统计，教育部门关于就学、毕业人口的统计，凡此种种，都可以做为直接来源数据的补充，同时也为评价直接来源的人口数据提供了可资对照的资料。

1.3 错误来源

任何来源的数据都不可避免含有错误，依据错误的严重程度，人口数据大致可分为三个等级。（1）完全不能使用，一些国家从未公布的普查便是一例。（2）有不同程度的缺陷，如使用则需对数据进行调整。（3）基本准确，可以直接进入分析。

产生错误的来源有哪些呢？首先是数据收集方式本身的缺陷，这种缺陷又称“结构性缺陷”，是由调查或统计方案的设计引起的。一个明显的例子是世界生育率调查，在这个调查的设计上，存在两个先天的限制，因为生育率调查是一种回顾性调查。