

经济和社会事务部  
统计司

方法研究

F辑 第92号

# 生育率和死亡率 数据收集手册



联合国  
纽约，2004年

联合国秘书处的经济和社会事务部(经社部)在经济、社会和环境领域的全球政策与国家行动之间起着重要的桥梁作用。该部的工作主要涉及三个相互关联的领域：(一) 汇编、制作和分析范围广泛的经济、社会和环境数据与信息，供联合国会员国在审查共同问题和评价政策抉择时加以使用；(二) 促进会员国在许多政府间机构内就采取什么联合行动方针对付现有的或新出现的全球挑战进行谈判；(三) 就将联合国各次会议和首脑会议上制定的政策框架转化为国家一级方案的方式方法向有关政府提供咨询意见，并且通过技术援助协助国家能力建设。

## 说 明

本出版物中使用的名称以及材料的编排方式，并不意味着联合国秘书处对任何国家、领土或其当局、或者对其边界或界线的划分表示任何意见。

联合国文件都用英文大写字母附加数字编号。

本报告中使用的“国家”一词在适当情况下亦指领土或地区。

提到商行和商业产品的名称并不意味着联合国赞成这些名称。

ST/ESA/STAT/SER.F/92
----------------------

联合国出版物
--------

出售品编号：C.03.XVII.11
--------------------

ISBN 92-1-561001-4

版权©联合国，2004年

版权所有

## 目 录

章	段 落	页 码
导言	1-23	1
A. 手册的目的	1-3	1
B. 数据收集方法	4-14	1
C. 实地调查工作的重要性	15-17	2
D. 有关出版物	18-23	2
一、制定生育率和死亡率数据的收集计划	24-79	5
导言	24-26	5
A. 生育率和死亡率数据	27-55	5
1. 民事登记、普查和调查记录	28-32	5
2. 出生、死亡和个人的点数	33-35	5
3. 比率和累加测度	36-40	6
4. 地理分类	41-48	6
5. 时间和时期	49-55	7
B. 审查过去的的数据收集活动	56-66	7
1. 民事登记	57-58	7
2. 人口普查	59-60	8
3. 一般户口调查	61-62	8
4. 生育史调查	63-64	8
5. 其他方法	65	9
6. 总体评估	66	9
C. 制定今后的数据收集活动计划	67-78	9
1. 民事登记	68-72	9
2. 人口普查	73-75	9
3. 一般户口调查	76	10
4. 生育史调查	77	10
5. 其他方法	78	10
D. 协调数据收集活动	79	10
二、实地工作、数据处理和存档	80-139	11
导言	80-82	11
A. 实地工作	83-104	11
1. 规划	88-90	11
2. 表格设计	91-94	12
3. 招募和培训	95-99	12
4. 督导	100-102	12
5. 信息的传输	103-104	13
B. 数据处理	105-126	13

章	段 落	页 码
1. 人工处理	107-110	13
2. 数据输入	111-113	13
3. 计算机编辑	114-122	13
4. 制表	123-125	14
5. 其他工作	126	14
C. 存档	127-139	15
1. 功能	130-133	15
2. 内容	134-136	15
3. 安全和维护	137-139	15
三、评价、估计和传播	140-226	16
导言	140-142	16
A. 评价	143-180	16
1. 活动的质量	145-146	16
2. 关联性和适时性	147-148	16
3. 数据质量	149-161	16
4. 估计的准确性	162-175	18
5. 误差的指示值	176-178	20
6. 抽样误差	179-180	20
B. 估计	181-193	20
1. 民事登记数据	183-186	20
2. 普查和调查数据	187-190	21
3. 根据多次估计值产生的水平和趋势	191-193	21
C. 传播	194-226	21
1. 印刷出版物	196	21
2. 数据集	197-203	22
3. 数字出版物	204-209	22
4. 数字格式	210-212	23
5. 数字复印	213-216	23
6. 标准通用置标语言 (SGML) 与可扩展标记语言 (XML)	217-226	23
四、作为生育率和死亡率数据来源的民事登记	227-379	26
导言	227-231	26
A. 民事登记	232-281	26
1. 中央民事登记机构	236-237	26
2. 地方登记处	238	27
3. 登记程序	239-243	27
4. 出生和死亡登记记录	244-248	27
5. 民间社会的了解和责任	249	28
6. 制表方案	250-271	28
7. 出版计划	272-275	30
8. 民事登记数据的质量	276-281	30

章	段 落	页 码
B. 利用民事登记数据计算出生率和死亡率	282-323	31
1. 承受风险	286-287	31
2. 按常住地址分类	288-292	31
3. 估计分母	293-308	32
4. 国家一级的出生率和死亡率	309-313	33
5. 国以下各级地理区域的出生率和死亡率	314-318	34
6. 人口亚群的出生率和死亡率	319-323	34
C. 登记完整性的概念	324-347	35
1. 未登记出生和死亡事件的组成成分	328-330	35
2. 选择性登记	331-334	35
3. 发生日期和登记日期	335-337	35
4. 延迟登记的分析	338-342	36
5. 登记完整性的趋势	343-347	36
D. 不完整民事登记数据的用途	348-363	37
1. 不完整登记的指示值和估计	352-355	37
2. 与普查和调查数据的总体比较	356-359	37
3. 利用模式	360-363	38
E. 人口登记作为生育率和死亡率数据来源	364-379	39
1. 概念	365-367	39
2. 登记迁入移民和迁出移民	368-372	39
3. 登记居住地址的变化	373-374	39
4. 登记其他个人特征的变化	375-376	40
5. 建立并保持人口登记册	377-379	40
五、来自普查和调查的生育率数据	380-559	41
导言	380-382	41
A. 存活倒推	383-419	41
1. 问题	386-391	41
2. 制表	392-393	42
3. 估计	394-397	42
4. 质量	398-406	42
5. 年龄分布误差	407-419	43
B. 亲生子女法	420-443	44
1. 问题	422-426	44
2. 制表	427	45
3. 估计	428	45
4. 配对	429-436	46
5. 质量	437-443	46
C. 生育史重建	444-472	47
1. 问题	449	47
2. 制表	450-462	47

章	段 落	页 码
3. 估计	463-471	49
4. 质量	472	51
D. 平均生育数	473-516	51
1. 问题	475-480	52
2. 制表	481-488	52
3. 估计	489-508	54
4. 质量	509-516	55
E. 最近出生人数	517-538	56
1. 问题	519-524	57
2. 制表	525-528	57
3. 估计	529-530	58
4. 质量	531-538	58
F. 生育史	539-559	59
1. 问题	542-550	59
2. 制表	551-552	60
3. 估计	553	60
4. 质量	554-559	60
六、人口普查和调查的死亡率数据	560-665	62
导言	560-564	62
A. 出生婴儿的存活情况	565-578	62
1. 问题	566-568	62
2. 制表	569-572	63
3. 估计	573-574	63
4. 质量	575-578	63
B. 根据生育史得出的死亡率数据	579-612	64
1. 问题	582	64
2. 制表	583-588	64
3. 估计	589-595	65
4. 质量	596-612	66
C. 近期住户死亡情况	613-635	70
1. 问题	618-623	71
2. 制表	624-625	72
3. 估计	626-627	73
4. 质量	628-635	73
D. 父母存活情况	636-651	74
1. 问题	637-638	74
2. 制表	639-640	75
3. 估计	641-643	75
4. 质量	644-651	75
E. 兄弟姐妹的存活状况	652-665	76
1. 问题	653-658	76

章	段落	页码
2. 制表	659	77
3. 估计	660	77
4. 质量	661-665	77
七、结论	666-681	79
A. 人口研究	667-668	79
B. 技术发展	669-670	79
C. 用户和用户需要	671-673	79
D. 数据收集环境	674-678	79
E. 国际合作	679-681	80
<b>附    件</b>		
列克西斯图	682-733	81
A. 列克西斯图表示法	683-691	81
B. 极端法	692-694	82
C. 交点法	695-702	82
D. 年龄别死亡率	703-707	84
E. 婴儿死亡率	708-712	85
F. 存活倒推法	713-718	85
G. 询问亲生子女数目方法	719-722	86
H. 平均生育数	723-726	87
I. 生育史	727-733	87
术语表		90
参考资料		94
表格		
表1. 人口数据来源的简要比较：目的、方法、人员配置和所需资源		3
表2. 人口数据来源的简要比较：提供信息的时间和特征		4
表3. 亲生子女表		45
表4. 生有一个或多个子女的妇女，按头胎的出生年份和头胎出生年份末妇女的年龄列示		48
表5. 生有两个或两个以上子女的妇女，按第二胎的出生年份和第二胎出生年份末妇女的年龄列示		49
表6. 生有两个或两个以上子女的妇女，按第二胎的出生年份和到第二胎出生那年年末妇女拥有1胎的整年数列示		50
表7. 生有三个或三个以上子女的妇女，按第二胎的出生年份和第三胎出生那年年末妇女拥有2胎的整年数列示		50
表8. 按年龄和平均生育数列示的妇女人数以及按妇女年龄组列示的平均生育总数		53
表9. 根据生育史获得的生育数		61

章	段 落	页 码
表10.	用于估计死亡率的出生婴儿和存活儿童表	63
表11.	按出生月份（调查访问前几个月）分列的儿童和按出生月份 （调查访问前几个月）和以周月表示的死亡年龄分列的死亡儿童	65
表12.	缅甸根据生育史数据得出的婴儿和儿童死亡率统计资料	68
图表		
图1.	巴基斯坦总和生育率估计值比较	19
图2.	根据生育史推算出的死亡率数据的列克西斯图示	66
图3.	说明死亡概率计算方法的列克西斯图	66
图4.	显示根据生育史数据计算的死亡率估计的年龄挑选择偏差的 列克西斯图	69
方框		
方框1.	数据、统计和人口：术语	8
方框2.	可扩展标记语言（XML）图示	24
方框3.	近期住户死亡情况的问题，1988年坦桑尼亚联合共和国人 口普查	71
方框4.	近期住户死亡情况的问题，1988年马拉维人口普查	72
方框5.	近期住户死亡情况的问题，1990年中国人口普查	72
方框6.	近期住户死亡情况的问题1991年缅甸人口变化和生育率调查	73
附件图表		
图A.1.	列克西斯图表示法	81
图A.2.	期间-同组人事件	83
图A.3.	年龄-同组人事件	83
图A.4.	年龄别死亡率	84
图A.5.	婴儿死亡率	84
图A.6.	存活倒推法	85
图A.7.	亲生子女法	87
图A.8.	平均生育数	88
图A.9.	生育史数据	89



## 导 言

### A. 手册的目的

1. 社会和经济的规划和监测需要有关出生、死亡和人口数据的资料。例如，为提供妇幼保健服务制定计划时要考虑到未来几年的预期出生人数。学校建造和教师培训将会受到不同年级预期学生人数的影响。制定就业政策时部分以加入和退出劳动力队伍的预期人数作为根据。社会保障制度必须预测即将领取养老金者的人数。

2. 关于出生、死亡和人口的数据通过几种互补的方法收集。在不同情况下可以对方法加以不同的组合。本手册详细介绍了现有的可用来确定什么样的组合最适合本国国情的方法。此外，它还介绍了如何利用通过这些方法编制的的数据推算出基本的生育率和死亡率指标。

3. 本手册无论对负责国家统计工作的政府官员和计划制定人员、国家统计局领导、户口登记和生命统计局主管还是对这些组织机构中从事专题事项和数据处理工作的专家来说，都是一项有用的工具；对于学者、研究人员和其他使用生育率和死亡率数据的人员来说也同样具有实用价值，因为了解了有关这些数据是如何产生的一些情况后，他们可以从中获益。本手册对于为统计工作提供资金和/或技术援助的本国和国际组织的工作人员也十分重要。

### B. 数据收集方法

4. 收集生育率和死亡率数据的方法通常有三种：人口普查、民事登记和户口调查。

- 人口普查收集国家范围内的每一个人和每一个住户在某个特定时点的经济和社会特征方面的资料。人口普查一般每10年进行一次。人口普查数据的一个重要用途是设计下文所述的人口调查的样本
- 民事登记收集有关发生在一国的出生、死亡和其他人口动态事件的信息。像人口普

查一样，民事登记的目的是全面覆盖。与人口普查不同的是，民事登记是一项持续性的行动。出生和死亡必须在事件发生后很短的时间内得以登记

- 户口调查收集关于相对较小但经过科学设计的住户样本的资料。与人口普查和民事登记相比，相对较小的样本规模可以降低调查费用，增加灵活性，但是削弱了提供有关较小地区和人口亚群的详尽信息的能力

5. 这三种方法中，没有一种是单独存在的或者甚至主要是为了收集生育率和死亡率的目的。例如，人口普查一般提供有关居住地点和移徙、住户和家庭特征以及教育和经济特征方面的信息，另外还提供人口方面的信息。进行民事登记的目的主要是为了建立有关出生、死亡和其他人口动态事件的法律文件。户口调查是用来收集许多不同类型的信息，极少仅限于收集生育率和死亡率数据。

6. 作为生育率和死亡率数据的来源，人口普查、民事登记和调查相互之间起到一种互补的作用。一旦完整的登记制度得以建立，民事登记便成为收集有关出生和死亡事件的基本数据以及有关死亡原因的数据的理想方法。一般说来，需要通过人口普查来提供计算生育率和死亡率以及累加测度所需的人数。人口普查数据还常常用来设计户口调查中使用的样本。反过来，户口调查是获取有关影响生育率和死亡率条件的具体数据的最适当方法。由于户口调查方法灵活，费用相对较低，所以通常是收集有关现时关注专题的数据的最佳方法。

7. 大多数国家都从使用这三种方法的过程中认识到这种互补性。关于收集生育率和死亡率数据的决定不是是否进行一次人口普查或户口调查，或者执行一项民事登记制度，而是如何对使用这三种方法收集生育率和死亡率数据进行最佳协调。

8. 民事登记是收集出生和死亡数据的最费力

的方法。世界人口中有一半以上居住在完全没有民事登记或提供的出生和死亡数据覆盖面不完整的国家。过去几十年中，提供完整的出生和死亡数据覆盖面的民事登记系统普遍取得了发展，在有些情况下甚至取得了更加长期的发展。民事登记系统的建立和初步发展的正确性通常是通过其初级法律功能而不是其中级统计功能得以证明的。

9. 在缺乏充分完善的民事登记系统的情况下，使用人口普查和户口调查来收集出生和死亡数据。虽然人口普查是一项艰巨而复杂的工作，但比起建立和维护一个充分完善的民事登记系统来则要省力得多。几乎世界各国都至少进行过一次人口普查，并开展过合理的彻底的人口清查。大多数国家都大概每隔10年左右时间进行一次普查，这种情况已经持续了数十年。

10. 人口普查和民事登记将提供地方一级的信息，甚至提供有关最小的地区的信息，因为它们覆盖所有的人和所有的出生和死亡事件。例如，它们还提供一国每一个初级卫生保健设施的受托区的人数，从中可以计算出服务提供的充分性的重要指标。初等和中等学校是为相对较小地区服务的设施的其他实例。

11. 然而，如果所需的资料只是关于整个国家或较大的地区，那么基于科学设计的样本的户口调查提供数据时所花费的费用远远低于人口普查或民事登记系统所需的费用。规模较小的户口调查行动具有使它们更加适合不断变化的信息需要的额外优点。

12. 少数国家利用人口登记加上一个充分完善的民事登记系统作为生育率和死亡率数据的来源。人口登记一般说来本身并不是一种数据收集方法，而是一种编制通过民事登记和其他方法收集而来的数据的方法。

13. 表1简要列示了作为生育率和死亡率数据来源的户口调查、人口普查、民事登记和人口登记的主要特征。完成各项原始资料收集工作所需的资源可比水平列于表格的最后两行。

14. 表2简要列示了四种来源提供的信息的时间和特点。两张表格的目的都在于给出对典型特征的高水平概述。有关详情见以下各章。

### C. 实地调查工作的重要性

15. 生育率和死亡率数据通过实地调查产生，在实地调查期间，一般公众的成员向代表数据收集组织的实地调查人员提供有关自身、其家庭成员及其所居住的住户的情况。通过任何一种方法收集的数据的完整性和准确性取决于实地调查工作的质量。在这个意义上，实地调查人员是每一次数据收集行动中的最重要的人员。不应该以他们处于组织系统的最底层为理由来掩盖这个基本事实。

16. 在数据收集工作中，没有哪一个方面比实地调查人员的招聘、培训和监督管理更加重要。产生统计资料是一个漫长而且往往十分复杂的过程。最后的结果不会比这个连锁过程中的最薄弱环节更好。实地调查是最马虎不得的环节。如果第一步努力失败，工作的随后几步有时可以重做。实地调查却不能重复，实地调查的失败的大部分后果无法通过后来的处理阶段得以消除。

17. 数据收集工作所受到的一个重要的实际限制是，随着问题的增加，并随着问题的复杂性的增加，所需要的培训和管理的广泛性也就增加，如果要得到有用的信息的话。人口普查和户口调查计划员常常被增加各种问题的请求弄得应接不暇。这些压力必须得到抵制，除非可以获得招聘、培训和管理实地调查人员所必要的资源。

### D. 有关出版物

18. 关于人口普查的论述见《人口和住房普查的原则和建议，第1次修订本》（联合国，1998年a）。另见《人口和住房普查的普查管理手册》（联合国，2000年a）、《人口和住房普查编辑手册》（联合国，2001年a）和《地理信息系统和数字化地图手册》（联合国，2000年b）。Cho（1976年）以及Cho和Hearn（1984年）的论著中列有许多亚洲和太平洋国家的普查调查表复印件。

19. 关于民事调查系统的论述见《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）以及五卷的《民事登记和生命统计系统手册》（联合国，1998年b、1998年c、1998年d、1998年e和1998年f）。

20. 《户口调查手册》（联合国，1984年）讨

论了户口调查问题。《综合户口调查方案的抽样基础和样本设计，初步版本》（联合国，1986年）、联合国调查表的编制和设计（1985年）、联合国的抽样误差（1993年）以及《户口调查中的抽样误差：资料来源、评估和控制，初步版本》（联合国，1982年）讨论了样本的设计和執行。

21. Cleland和Scott的论著(1987年)回顾了世界生育率调查方案的经验，还提供了阅读具体旨在了解生育率调查的文献资料的指导。《1991年8月5日-7日人口和健康调查世界大会》对人口和健康调查方案的经验进行了审查(资源开发/宏观国际研究所，哥伦比亚特区华盛顿，1991年)。各国国别报

告以及世界生育率调查方案和人口和健康调查方案的有关出版物中也作了大量详尽介绍。

22. 《人口登记和类似系统的方法和评价》(联合国，1969年)中也讨论了人口登记问题。只有少数国家把人口登记作为生育率和死亡率数据的来源。

23. 美国国家科学院于1970年代末进行了一次关于生育率和死亡率估计数据收集的审查(国家科学院，1981年)。另见Cleland的调查论文(1996年)。Casley和Lury的论著(1981年)中对发展中国家的数据收集作了非常一般性的处理。

表1. 人口数据来源的简要比较：目的、方法、人员配置和所需资源

项 目	数据来源			
	户口调查	人口普查	民事登记	人口登记
目的	与全国代表性住户样本中的人员面谈	在特定时点点查该国的每一个住户和每一个人员	持续登记发生在该国的每一件出生和死亡事件	保持每个人的持续更新记录
方法	向每一个样本住户连续几个月收集信息(单轮调查一次，多轮调查反复几次)	在特定时点向该国的每一个住户收集信息	建立和维护一个永久性的全国范围的地方登记处网络，登记处工作人员由地方登记员组成，他们负责收集和记录关于生命动态事件的信息	根据出生、死亡、迁入移民、迁出移民、国内迁移和个人特征变化建立并不断更新人员记录
工作人员配置	参与培训和实地调查期的访问员和监督员人数应当足以覆盖样本地区	参与培训和实地调查期的点查员和监督员人数应当足以覆盖整个国家	全国各地长期在地方登记处工作的地方登记员	操作和维护登记系统的工作人员，负责登记出生、死亡、入境移民、出境移民、居住地址变化以及其他个人特征变化
所需资源的可比水平	所需资源一般少于其他任何方法	由于业务量巨大，所需资源大于调查	所需资源大于普查，原因是需要有永久性的、全国范围内的设施和工作人员	所需资源大于民事登记，因为必须登记国内和国际移民以及个人特征变化

表2. 人口数据来源的简要比较：提供信息的时间和特征

项 目	数据来源			
	户口调查	人口普查	民事登记	人口登记
数据收集的时间	灵活机动	通常10年一次	持续性	持续性
报告的时间	一般实地调查后1个月至1年	完整的点查结果一般1-3年	一般是报告年度结束后一年内的年度报告	经验不足
数据时间基准	实地调查前15年或15年以上的数据	基准日期前15或15年以上的数据（成人死亡率只需1-2年的数据）	灵活（受登记延迟所造成的限制的影响）	灵活
覆盖地理范围	仅仅包括国家和主要民政区（样本范围越大，提供的情况越详尽）	所有地理区域	所有地理区域，但受不完整登记的影响	所有地理区域，必须准确登记所有迁徙情况
其他信息详情	有可能很详尽，但不包括小群体、不常发生的事件或死亡原因	基本信息	基本信息加上死亡原因	基本信息加上死亡原因
灵活性	最灵活	灵活性不如调查；普查规模所需的问题很少，相对比较简单	灵活性不如普查；难以改变关于登记方式的项目	灵活性不如民事登记；难以改变获取个人特征变化数据的系统

## 一、 制定生育率和死亡率数据的收集计划

### 导 言

24. 生育率和死亡率数据来源通常不止一个。例如，当出生和死亡数据获自民事登记时，计算比率和累加测度所需的相应人数通常根据人口普查数据估计。如果利用人口普查收集出生和死亡数据，它们常常通过各种调查得以补充，因为各种类型的调查可以提供更加具体和更加及时的数据。不论利用哪一种数据来源，对各种来源的数据进行评估时通常包括与其他来源的数据进行比较。

25. 因此制定生育率和死亡率数据的收集计划包括两个不同的阶段。第一阶段确定将从哪一种来源获取什么样的生育率和死亡率数据，同时处理不同来源之间的协调问题。第二阶段把这种信息纳入民事登记系统的规划、下一次人口普查、即将进行的户口调查或者涉及的任何一种数据收集行动。

26. 民事登记系统、人口普查和户口调查的规划在有关这些方法的文献资料中作了广泛讨论。本章重点讨论综合这三种来源的数据资料所需要的规划问题。但是在处理这些问题之前，必须具体界定几种不同类型的生育率和死亡率数据。关于这一问题在下一节中进行讨论。

#### A. 生育率和死亡率数据

27. 有几种不同类型的生育率和死亡率数据，范围从有关特定出生和死亡事件的信息一直到整个国家的生育率或死亡率水平的总指标。A节的讨论范围包括各种不同类型的数据及其相互之间的关系。

##### 1. 民事登记、普查和调查记录

28. 从最基本一级讲，生育率和死亡率数据包括有关特定出生、死亡事件和个人的信息。

29. 民事登记系统提供关于特定出生和死亡事件的数据资料。必须区分事件的法律记录与统计

记录，这一点很重要。出生和死亡事件的法律记录是民事登记的首要输出信息。建立统计记录的目的是汇编出生和死亡数据，关于这一点将在下一节作进一步研究。

30. 出生和死亡事件的统计记录包括有关事件的数据资料，例如事件的发生日期和地点；有关经历事件的人的数据资料，例如死亡时年龄、出生时体重和出生时母亲的年龄；以及有关事件登记的资料，例如登记的日期和地点。有关其他详尽论述，见第四章A节第4小节和《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）。

31. 人口普查产生人口中的住户和个人在一个特定时点的记录。个人记录包括诸如居住地点、性别和出生日期或普查时年龄等信息。人口普查提供的有关个人的信息一般不如户口调查那么具体，但是提供有关人口中每一个人的信息，而不仅仅提供有关个人样本的信息。关于建议列入人口普查的主题清单，见《人口和住房普查的原则和建议，第1次修订本》（联合国，1998年a）。

32. 户口调查也产生有关人口中的住户和个人的记录，但是数据资料通常提到调查日期，而不提固定的基准时间。户口调查中包括的个人数据资料通常比普查更为详尽，而且只涉及人口的一个样本。

##### 2. 出生、死亡和个人的点数

33. 为了统计说明和分析的目的，关于各人出生、死亡和个人的信息资料用不同方法加以归纳。最基本的汇总数据是不同群体中的出生、死亡和个人的点数。

34. 在民事登记的术语中，汇编出生记录的目的是为了产生在某个历年期间或其他时期的总的出生人数，出生事件按出生婴儿性别、婴儿出生时母亲年龄和出生婴儿、母亲或有关人员

的其他特征分类。同样，汇编死亡记录的目的是为了产生死亡的人数和分布。关于出生和死亡的建议制表目录，见《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）。

35. 在普查和调查的术语中，对个人记录进行制表的目的是为了产生按性别、年龄和其他特征分列的总人数和分布。有关建议的普查制表的目录，见《人口和住房普查的原则和建议，第1次修订本》（联合国，1998年a）。

### 3. 比率和累加测度

36. 用一个代表处于经历这些事件可能的人的相应数字除一个出生和死亡人数，得出的相对数字即为出生率和死亡率。出生和死亡数单独并不传递有关生育率和死亡率水平的有用信息，因为它们反映人口的规模和观察的时间长度。例如，在中国得到的出生率预计比摩纳哥高，因为中国的人口高于摩纳哥，预计10年期间的出生率高于单独一年中的数字。

37. 用于生育率研究的最重要的比率是年龄别出生率，就是各个年龄组的生育数与这些年龄组的女性人数的比率。用于死亡率研究的最重要比率是年龄别死亡率，就是各年龄组的死亡人数与这些年龄组的人数的比率。年龄别死亡率始终应当男女分列，但也可以把男女性别放在一起列示。

38. 出生率和死亡率可以按非年龄特征或者除年龄以外的其他特征分类。例如，出生率可以按年龄和一名妇女生育的子女数分类，死亡率可以按年龄和死亡原因分类。

39. 对按年龄和/或其他特征分类的一组比率中所载的数据资料进行归纳时，通常采用从这些比率中计算出一个或多个累加测度的办法。例如，生育水平的一个基本指标是总和生育率，其计算方法是把所有育龄妇女的年龄别出生率加在一起。总和生育率可以解释为一名到育龄期结束时存活的妇女一生中预期生育的子女数，如果她经历了给定的年龄别出生率的话。

40. 死亡率水平的基本指标是婴儿死亡率和出生预期寿命。婴儿死亡率表示预期在一周岁以前可能死亡的出生婴儿所占的比例。出生预期生命

表示刚出生的婴儿预期可能存活多久，如果这名儿童经历了在一个给定年份或其他时期所观察到的年龄别死亡率的话。预期寿命是众多累加测度中的一个，可从通过年龄别死亡率计算出来的生命表推算。

### 4. 地理分类

41. 各种人口数据涉及定位于空间和时间的个人、出生和死亡。这一小节讨论地区（空间）分类，下一小节讨论时间和序时分组问题。

42. 从人口普查和民事登记中推算生育率和死亡率数据时，制表计划的最困难的方面之一涉及将要提供的地区详情。就调查而言，地区详情引起的问题要少一些，因为抽样大大减少了这种可能性。

43. （个人）居住“地点”或（出生和死亡）事件只有在关系到一些地区分类系统时才具有意义。最常见的分类系统是城市和农村地区、大小行政区以及主要城市和城镇。

44. 事实上，制表时普查数据可以低到普查点数区一级，民事登记数据可以低到初级和中级登记单位一级。但是，这种单位可能多达几千甚至几万个，对较大的国家来说，可能会有好几百万个这样的单位。编制数量浩大的所有单位的详尽表格通常既无用处也不可行。

45. 因此必须确定哪一种制表将包括哪种级别的地区详情。由于可能为任何单独一个地区编制几十个或几百个表格，加之地区数量十分巨大的地区分类，很难作出这种决定。在利用制表计算各种出生率和死亡率时，要进一步考虑普查与民事登记表之间的协调。

46. 一项普遍原则是国家一级提供的表格将会越来越详尽，大行政区、小行政区和更加详尽的地区分类提供的数据表将会陆续减少，而且越来越不详尽。虽然这项原则就其本身而言是有用的，但在对哪一种地区分类应当编制哪一种表格的问题上，它没有提供任何指导。

47. 事实上，普查和民事登记的全面覆盖率使得能够为除了行政单位的常见嵌套分层以外的许多

不同种类的地区系统提供数据。可以通过为尽可能最小的制表单位综合数据的办法使所有各种地理区域尽可能地接近。可以为根据地形界定的区域编制数据，例如水源保护区、森林保留地或河流和海岸线邻近地区。

48. 最近几十年来地理信息系统的发展极大地便利了这种空间数据的编制和使用。《地理信息系统和数字化地图手册》（联合国，2000年b）对于一些特别涉及到人口普查的主题作了非常有用的介绍，但是许多信息同样与民事登记数据有关。

## 5. 时间和时期

49. 各种人口数据涉及到存在过一段时间的个人和在特定时点发生在这些人身上的事件。数据涉及的时间或时期应当始终明确。

50. 从定义上讲，人口普查是在某个时点点查人口，这个时点叫作普查的基准时间。例如，基准时间可以是2000年6月30日的午夜时分。实际上，一般确定普查的日期就足够了。

51. 民事登记数据涉及发生在某一特定时期内、日历年和日历月内、有时甚至更短的时期内的出生和死亡事件。日历期间从第一天的午夜开始到最后一天的午夜为止。基准时间在一个充分完善的民事登记系统中可以是非常准确的。

52. 户口调查的基准时间比较复杂。一些调查遵循获取截止某个具体时点的所有信息的普查规则，但是大多数调查获取截止调查时间的信息。由于实地调查可以延长好几个月，所以向不同住户收集的数据涉及不同的时点。因此人口数字严格说来比不上普查数据，出生和死亡数据严格说来也比不上民事登记数据。

53. 实际上，调查数据通常可以被认为好像涉及到一个特定时点，当作进行调查访问时间的中心趋势的某种测度。最简单的这种测度是进行调查访问期间的中点，对大多数目的来说这一测度就足够了。然而，常常是调查访问期较长，也许是6个月，大多数调查访问集中在调查访问期的开始。在这种情况下，调查访问的中间日期可以作为一个比较适当的基准时间。调查报告可以提供进行调查访问的这个月的调查分布情况，从中能够估计出这个

中间日期。

54. 人口统计资料正像从中得出的这些数据一样，始终有一个应当明确的基准时间。一般说来，提供的人的数字、百分比和比率是某一时点的；出生和死亡的数字、百分比和比率是关于某一个时期的；各种比率是针对某一个时期的。就户口调查数据而言，由于前面两段中所论述的原因，基准时间可能是不明确的，但这一般没有任何实际结果。

55. 某些情况下，比率涉及到时期的一般规则也会发生例外。例如，当婴儿和儿童死亡率用出生并存活儿童人数估计时，一些估计程序就产生涉及时点而不是涉及时期的比率（见第六章A节）。可以通过利用时期的中点确定这些时期的比率并在两个中点之间插入各个时间的比率，将这些比率与涉及时期的比率进行比较。

## B. 审查过去的的数据收集活动

56. 制定生育率和死亡率数据收集计划时应首先审查现有数据资源，目的是归纳已经获得的是什么样的资料，并评估现有资源满足用户需要的程度。这种审查主要将包括对关于各种现有数据来源的一系列问题进行解答。以下各小节简述有关各种来源可能提出的问题。

### 1. 民事登记

57. 在出生和死亡事件的登记达到完全覆盖的情况下，民事登记是基本生育率和死亡率数据以及死亡原因数据的较理想来源。在民事登记没有达到完全覆盖，但是获得的出生和死亡事件的数据占据了相当大比例的情况下，它可能是生育率和死亡率数据的一个宝贵来源。关于民事登记作为生育率和死亡率数据来源的讨论，见第四章。

58. 与民事登记有关的问题包括如下：是否有国家民事登记系统？它是否编制按年龄和其他相关特征分列的年度出生和死亡表？如果是的话，它的有效期多长？出生登记的估计完整性有多大？死亡登记的估计完整性有多大？年龄别出生率和死亡率估计数和其他统计数字是从这些数据中产生的吗？如果是的话，这些估计数的满意程度有多大？

使用什么方法得出这些结论的？

### 3. 一般户口调查

#### 2. 人口普查

59. 人口普查是生育率和死亡率数据的一个可能的丰富来源。像民事登记一样，人口普查提供有关所有地理区域的数据，而且它还可提供关于可从普查问题角度界定的任何人口亚群的数据，例如受教育程度、职业或移徙状况。民事登记即使在充分完善的情况下，提供的数据常常也不会像人口普查那么详尽。追溯性问题和方法常常允许编制出普查日期前15或15年以上的生育率和死亡率数据。从人口普查中产生生育率和死亡率数据的问题和方法在第五章（生育率）和第六章（死亡率）中讨论。

60. 关于人口普查的问题的实例包括如下：最近一次人口普查是什么时候进行的？任何关于出生和死亡的追溯性问题是否包括在内并用于估计生育率和/或死亡率？是否使用了亲生子女法产生普查前15年年龄别出生率的估计数？向整个人口还是仅仅向一个样本提出了这些问题？所得出的估计数的满意度有多大？使用什么评估方法得出了这些结论？

61. 一般户口调查也是生育率和死亡率数据的一个可能的丰富来源。一般户口调查不会提供地理详情或关于小规模人口亚群的情况，后者由普查提供，但是这种调查可能包括更加详尽的生育率和死亡率方面的问题，进行的次数可能比人口普查更加频繁。从一般户口调查中获得生育率和死亡率数据的问题和方法在第五章（生育率）和第六章（死亡率）中讨论。

62. 与一般户口调查有关的问题可包括如下：最近一次人口普查是什么时候进行的？任何关于出生和死亡的追溯性问题是否包括在内并用于估计生育率和/或死亡率？是否使用了亲生子女法产生普查前15年年龄别出生率的估计数？样本规模多大？样本设计如何？使用什么方法计算抽样误差？理想的生育率和死亡率统计的估计抽样误差为多少？

#### 4. 生育史调查

63. 生育史调查得出的生育率资料可能比其他任何来源更加详尽，但是涉及的广泛的问题通常

#### 方框1. 数据、统计和人口：术语

**数据**可定义为有关某种统计总体中的实体的系统资料。系统在这里意味着，只要失缺不适用于特定实体的数值和资料，就向统计总体中的每一个实体提供相同的资料。

**统计总体**的概念是很通用的。它只要求组成统计总体的各个实体互不关联，要求界定统计总体的成员资格规则明确规定给定类型的实体中哪些是成员，哪些不是成员。

统计总体的实例有：给定年份期间在某个人口中发生的出生或死亡事件；特定时点在某个人口中现存的住宅单元、住户和人员；以及样本住户和人员，其有关资料在户口调查中获取。一个国家的大大小的民政区和世界各国也是统计总体。

**统计量**是根据数据计算出来的一个数值。出生、死亡和人员以及各种比率和累加测度都是统计数字。统计数字，例如总和生育率或婴儿死亡率，常常是针对世界某些选定国家或某个特定国家的大小行政区收集的。从上一段的意义上讲，这种信息就是数据，其中统计总体由一些国家或一国的一些地区组成。因此“数据”是一个比它乍看起来更加包罗万象的概念。

“数据”和“统计量”在人口学和统计学中具有相同的意义，但是人口学家和统计学家使用“人口”一词时采取了完全不同的方式。在人口学中，人口是一个集合体，由于进入和离开人口的结果，其成员人数随着时间而变化。这就是将这个名词用于本手册时的预期意义。统计学家采用“人口”这个名词来描述数据所涉及的统计总体（Stuart和Ord，第1卷，1987年）。



使得它们局限于相对较小的样本。生育史调查一般包括关于有关主题的非常详尽的问题，例如影响生育水平和趋势的因素、节育措施的使用或关于儿童和家庭保健的情况以及卫生保健服务的可获得情况。生育史问题讨论见第五章F节。

64. 关于生育史调查可以提出诸如以下一些问题：最近一次生育史调查是什么时候进行的？调查是否包括完整的生育史？得出的生育率估计数的满意度有多大？是否存在任何误报出生日期而引起的偏差的迹象？是用什么评估方法得出这些结论的？样本规模多大？是否对理想的生育率和死亡率统计的抽样误差作出了估计？如果是的话，抽样误差多大？

## 5. 其他方法

65. 使用过其他方法吗，例如多次调查或双重记录系统？如果是的话，得出的年龄别出生率和死亡率估计的满意度多大？是用什么评估方法得出这些结论的？

## 6. 总体评估

66. 进行总体评估时的相关问题包括如下：可获得的生育率和死亡率数据的满意度多大？它们满足用户需要达到什么程度？如果存在的话，哪一种用户需要没有得到满足？可获得的数据提供的关于国以下各级地区和人口亚群的情况是否足够详尽？它们提供的详情是否足以描述趋势？各种相互一致的方法编制同类或类似的统计估计数时达到什么程度？认为估计数的精确度有多高？是用什么评估方法得出这些结论的？第三章A节概括性地讨论了评估生育率和死亡率数据的问题。关于民事登记数据的完整性估计的讨论见第四章D节。

### C. 制定今后的数据收集活动计划

67. 对当前活动进行审查时可以把重点放在众多过去没有加以利用的可能性上。C节简要说明了主要数据收集方法所提供的各种可能性。

## 1. 民事登记

68. 关于作为生育率和死亡率数据来源的民

事登记的今后活动可以处理以下一些事项：(a) 评价出生和死亡登记的覆盖率和选择性；(b) 提高民事登记的覆盖率；(c) 改进民事登记数据的汇编和出版；或(d) 改进对民事登记数据的利用。

69. 评估有关生命事件的资料的完整性和准确性是利用民事登记作为统计资料一种来源的一个重要的方面。第四章D节讨论了对出生和死亡登记的完整性的估计问题。如果最近没有作过这种评估，不妨着手进行一次。

70. 努力提高不完整民事登记的覆盖率是重要的，但是证明这种努力是否正确的通常是民事登记的初级法律功能以及它的中级统计功能。实现全面覆盖一般是一项长期目标。

71. 在民事登记系统已经建立但不产生年度的出生和死亡数字的情况下，相关的问题包括如下：什么问题正在阻止产生这些数字？在组织生命记录的处理方面必须作出什么样的改进以提供理想的统计结果？

72. 不完整民事登记数据很少得到尽可能充分的利用。即使获得的数据在全部出生和死亡数据中只占不大的一部分，这样的系统也可提供有关生育率和死亡率的宝贵数据。完整性的估计水平可用来调整出生和死亡的登记数字，以对总数作出估计，并适当考虑登记的选择性。

## 2. 人口普查

73. 今后的人口普查应当始终被看作是生育率和死亡率数据的一种可能来源。在民事登记数据不完整或不存在的条件下，这一点特别重要。需要提出的特殊问题极少，但是所提供的生育率和死亡率数据的价值可能很大。

74. 在出生登记没有捕获到全部出生事件的情况下，应当优先考虑采用亲生子女法提供普查前15年的年龄别出生率的估计数。进一步详情见第五章。

75. 在死亡登记没有捕获到全部死亡事件的数据的情况下，应当考虑将有关最近发生的死亡事件的问题列入用于估计成人死亡率水平目的的普查调查表。进一步详情见第六章。

### 3. 一般户口调查

76. 今后的一般户口调查应当始终被看作是生育率和死亡率数据的可能来源。在民事登记数据不完整或不存在的条件下以及在人口普查没有被用作生育率和死亡率数据的来源的情况下，这一点尤其重要。第五章（生育率）和第六章（死亡率）对问题和方法进行了探讨。

### 4. 生育史调查

77. 生育史调查的目的部分是为了提供大体上标准的生育率和死亡率资料。在制定生育史调查计划时，应当为推算生育率和死亡率统计例如总和生育率和婴幼儿死亡率的抽样误差的计算作出安排。另外，还应当对从这些可能提供最准确数据的调查中估计的生育率和死亡率水平和趋势与从其他来源得出的估计作系统比较，这一点也同样重要。对生育史问题的讨论见第五章F节。

### 5. 其他方法

78. 多次调查和双重系统法比单次调查复杂得多，而且费用也高昂得多，因此这种方法不应当成为收集生育率和死亡率数据的方法，除非利用其他方法收集数据的一致努力未能获得成功。如果为了收集其他种类的数据而必须进行多次调查的话，当然应当考虑将关于收集生育率和死亡率数据的适当问题包括在内。

#### D. 协调数据收集活动

79. 既然生育率和死亡率数据是从大量不同来源的数据中推算出来的，那么从任何一种来源产生数据的计划都应当考虑到从其他来源产生数据的计划。例如，当民事登记没有提供出生和死亡事件的完整报告时，利用人口普查和户口调查作为生育率和死亡率数据的来源就较为重要了。同样，如果普查没有被用于这一目的的话，那么利用一般户口调查产生生育率和死亡率数据就更显重要。

## 二、 实地工作、数据处理和存档

### 导 言

80. 生育率和死亡率数据的每一种收集方法都涉及实地工作，即向一般人口中的个人收集相关信息，以及数据处理，即将信息汇总后进行处理，从而提供更加有价值的统计信息。本章的前两节对这两种活动进行了综合说明，并着重强调与生育率和死亡率数据收集相关的内容。

81. 现在，数据处理工作几乎无一例外地由计算机承担，凭借大容量的存储器，现代计算机甚至可以保存最大的数据集，方便日后的重复使用。计算机越来越多地应用于规划、管理和行政工作，因此各种类型的业务文件都可以进行存档，并在日后的规划、评估和其他活动中用作参考。

82. 由于重心由打印文件和出版物转移到计算机文档，存档工作日益重要和艰巨。本章第三节和最后一节将会讨论存档问题。尽管这些问题都非常广泛，但是它们对于生育率和死亡率数据有着特别重要的关联性，因为生育率和死亡率数据在很多情况下都有着不同的数据来源。

#### A. 实地工作

83. 实地工作是“实地”收集有关人员和事件的信息的过程——全国性的民事登记或者人口普查，或者抽样地区的户口调查。要求公众向代表数据收集组织的实地工作人员提供有关本人、家庭和住户的信息。

84. 在民事登记中，申报人前往地方登记处进行出生和死亡登记，由地方登记员负责收集信息。第五章的A节将对登记过程作进一步的说明。

85. 在普查和调查中，信息收集方式通常是由点查员或者调查访问员走访全国各地的住户，从每户一名或多名被调查者那里获取信息。普查和调查也可以采用“自我点查”法，将表格邮寄给住户，并希望户主或者其他被调查人填写表格并通过

邮件寄回。对于没有填写表格并寄回的住户，才动用点查员和访问员。

86. 实地工作对数据收集组织的组织和后勤能力提出了很高的要求。大国的民事登记和人口普查需要成千上万的实地工作人员在全国各地开展工作。人员招募到位后，要经过培训、装备，最后还要接受督导。在普查和调查中，由于时间相对较短，而任务涉及全国各地，因此必须实施严密的协调和控制。

87. 特别是，收集填写完备的表格必须有周到的后勤程序，确保从现场运往数据处理地点的过程中表格不会发生损坏或者遗失。即使在较小的国家，人口普查也会涉及到几千万份表格，所有的表格都必须运送到一个或多个处理中心，在妥善存储后等待进一步的处理。

#### 1. 规划

88. 实地工作的规划必须考虑(a) 用于收集信息的表格的设计；(b) 实际开展工作的人员的招募和培训；(c) 提供开展工作所需的设备、必需品、参考和培训资料以及其他资源；(d) 监督工作人员的工作，找出问题并实施补救措施；(e) 监督信息的储存和传输，确保指定的数据处理中心及时完好地接收到信息。

89. 《人口和住房普查的原则和建议：第1次修订本》（联合国；1998年a）和《人口和住房普查的普查管理手册》（联合国；2000年a）中讨论了人口普查的规划问题。民事登记信息统计处理的规划问题也在《关于生命统计系统的原则和建议：第2次修订本》（联合国，2001年b）一书中有所述及。《户口调查手册》（联合国，1984年）讨论了户口调查的规划问题。

90. 关于生育率和死亡率数据的收集问题，规划应当事先考虑统计中需要哪些信息，哪些表格将

用来收集这类信息。至于普查和调查中收集的生育率和死亡率数据，规划应当考虑问题的事先测试以及对实地工作人员进行提问培训。测试在民事登记中可能也会发挥作用。

## 2. 表格设计

91. 无论民事登记、人口普查还是户口调查，统计表格的设计都是实地工作得以成功开展的基础。尽管首先考虑的问题是收集具体的信息，但是问题的措辞和表格格式可能会极大地促进或者阻碍完整而准确的信息收集。

92. 在人口普查和户口调查中收集有关出生和死亡的回溯性信息时，尤其需要注重表格的设计。出生和死亡在相关人员的生活中都是大事，很难被遗忘。但是，为了准确地获取相关信息，必须针对特殊的障碍，设计特别的问题序列。例如，人们可能不愿提起死去的亲人，或者出生后不久即夭折的子女。另外，有人尽管清楚记得事件发生了，但是却不确定发生的确切时间。

93. 尽管熟悉本国国情并通晓其他国家的经验会为表格的设计提供宝贵的背景资料，但是这些优势并不能替代实地测试。在确定哪种类型的问题能够产生最完整、最准确的信息报告时，务必将实地测试视作“最后的诉求手段”。

94. 第五章（生育率）和第六章（死亡率）详细述及人口普查和户口调查中的生育率和死亡率数据收集问题。

## 3. 招募和培训

95. 实地工作最重要的一方面就是实地工作人员的招募、培训和督导。这适用于地方民事登记员、点查员以及调查访问员。培训的双重作用在于，它既能向实地工作人员灌输工作中需要用到的具体知识，同时还能营造集体意识，工作人员可能会经常面对困难，集体意识将会鼓励他们在困难的条件下也尽可能地做好工作。招募的作用在于选拔最适合开展这项工作的实地工作人员。

96. 准确的信息收集要求实地工作人员细致深入地理解他们将要填写的表格上的每一项内容，要求他们了解填写表格的适当程序，了解碰到困难

以后如何寻求帮助。在普查和调查当中，点查员和访问员必须了解如何走进指定点查的住户家中。

97. 透彻、深入地理解收集上来的各项信息十分重要。表面看来简单的问题常常可能产生歧义，必须在培训中加以阐释。例如，有关曾平均生育数问题，子女指亲生子女，而不是养子养女；包括死亡以及尚存的子女；包括住在别处的子女以及和母亲住在一起的子女。

98. 在培训手册上应当收录对于所使用表格上各个项目的详细注释。在培训期间，每个实地工作人员都会领到一本手册，并在实地工作期间用作参考。

99. 包含更多问题和更多复杂问题的普查和调查，如果想要产生有价值的信息，就需要进行更加广泛的培训和督导。现实的问题是，人口普查和抽样调查活动常常面临压力，需要增加更多更复杂的问题，而工作中可以利用的资源却十分有限。如果实地工作人员没有充分接受有关问题的培训，产生的问卷信息可能没有太多价值或者根本没有价值。

## 4. 督导

100. 对实地工作人员的督导是保证工作成功开展的关键。花在培训上的时间和资源总是有限的，即使接受了最好的培训，随着工作的开展，培训的部分内容也会被遗忘。实地工作中也会涌现出一些培训时未能预料的问题或困难。在开展实地工作的过程中，对工作人员进行督导，是保证工作人员尽可能做好工作并从被调查者那里获取准确信息的必要手段。

101. 督导人员本身也是实地工作人员，必须接受任务培训，另外对他们也必须进行适当的监督。实地工作人员和他们的督导员必须定期会面，讨论出现的问题以及解决问题的方法。在实地普查和调查工作中，督导员应当每天与查点员会晤，回顾一天工作当中出现的问题，并探讨解决这些问题的途径。

102. 有效督导的另外一个要素是通过实地编辑的方式，监督实地工作人员个人的任务完成情况。督导员审查某实地工作人员制作的表格是否有

错误、遗漏或者含糊不清的地方。一旦发现任何问题，当即与制作表格的实地工作人员开展讨论。尽量做到修正错误、拾遗补漏并澄清意思含糊的地方。不过实地编辑最主要的目的还不是改正错误，而是防止以后再次发生错误。督导员时刻掌握各实地工作人员的工作质量，并及时地解决出现的疑惑、错误和疏漏，防止其影响实地工作的顺利开展。有关这一问题的进一步讨论，见《人口和住房普查的普查管理手册》（联合国，2000年a）。

## 5. 信息的传输

103. 传统上通过纸张形式的表格、“图表”或“问卷”来记录信息。随后这些表格将会被运送到一个或多个中心地点，接受进一步的处理。将会采用不同的程序，确保表格不发生丢失或者分类错误。

104. 现在，大多数情况下依然使用纸张表格来收集信息，但是有时也会将信息直接输入到计算机中。如果采取直接的计算机输入，可以将信息传送到适当的计算机储存媒体中，例如软盘或者CD-ROM，或者通过电话线或计算机网络，传输到中心地点，进行进一步处理。

## B. 数据处理

105. 在数据处理中，输入数据是从被调查者（普查和调查）或者申报人（民事登记）那里实地收集的信息，而产生的输出数据是数据集、制表以及其他导出统计数据，以及关联文件和数据处理报告。

106. 民事登记系统、人口普查和户口调查工作中的数据处理涉及同样的基本操作。民事登记系统的数据处理为持续性操作，或者是定期性操作但间隔相对较为频繁，而普查或者调查的数据处理时间通常是实地工作完成后的一段固定时期。

### 1. 人工处理

107. 进行数据处理时，首先应当通过一定的程序，核查已经收讫所有实地收集的信息并且信息的形式适于进一步的处理。还将会核查记录，确保其中包括基本的身份识别信息。

108. 在民事登记中，为统计目的进行数据处理的负责机构通常会定期收到有关出生和死亡的统计报告。对于是否即将在规定时间内收取统计报表予以核实，并针对表格的内容进行一些基本核查，确保可以进行数据的输入。

109. 在人口普查中，将根据每个普查区的汇总表，核对从各区收上来的调查问卷份数，确保实地收集的全部表格如数到达处理地点。普查通常收集住户和人员的信息，包括该住户的成员数。

110. 有些信息，例如职业或者行业的普查或调查信息，可能在数据输入之前需要进行人工编码。不过，随着信息技术的发展，计算机辅助编码和自动编码可能会取代人工编码。下文将会对此进行详细的说明。

## 2. 数据输入

111. 各种数据收集活动一般旨在及早将信息传送到信息储存媒介，从而尽量减少人工处理环节。如果信息收集于纸张表格中，可以将表格中的信息键入计算机或者从若干光扫描技术中选用一种，完成数据输入，或者“捕获”。

112. 无论使用哪一种方法，都必须将生成的计算机记录信息与原始表格上的信息进行系统对照，核实输入过程的准确性。通常采用抽样方式进行核对。如果抽样显示出问题，则应对该组记录进行全面核对。

113. 如果信息直接输入地方登记处的电脑，或者在实地普查或调查时即已完成输入，数据输入便不再是独立的操作。将口头答卷转换为计算机编码的工作由计算机软件来完成。一旦确立了软件的精确度，其他数据输入方法所需要的核对程序就显得多余的了。随着所需的计算机硬件成本下降，今后在实地工作期间直接将信息输入计算机可能会变得日益普及。

## 3. 计算机编辑

114. 输入数据并进行必要的分类操作后，所产生的数据集需要进行一系列由计算机执行的编辑。每次编辑包括编辑检查，如果测试失败，则进行相关的编辑操作。

115. 结构检查测试特定记录存在与否。例如，可以测试任何数据集中的全套记录，确保每个地方登记区或查点区都得到反映。这种测试通常与上文B节第1小节所描述的人工检查同时进行。如果缺少任何地区的记录，则会找到相关记录，并纳入数据集中。

116. 可能进行其他结构检查，确定是否每一住户只有一份户主/被调查人的记录，是否在数据输入时错误地将同一住户的记录复制了两份，或者每户成员人数的记录是否与户口记录上显示的成员数目一致。

117. 将会检查单份记录，核实是否存有无效代号和漏测值。无效代号通常表明数据输入时出现错误，可以查找原始的纸张表格予以修正。无效代号检查常常纳入数据输入核对过程。

118. 遗漏的信息最好从信息来源处查找，即包含该记录的纸张表格，或者最初提供该信息的申报人或者被调查者。如果无法查找信息来源，应当使用“不确定”编码，数据表显示对应的类别，例如，“年龄不确定”。应当始终使用明确的“不确定”代号，避免混淆遗漏的信息和不适用的信息。应当认真确保漏缺值代号和值不发生混淆。

119. 不完整数据有时是估算数，即在统计上合理但未必正确的指定数值。估算不宜多用，因为提供不完整数据虽然方便用户，但不当的估算可能会破坏数据。

120. 例如，对于平均生育子女数和存活子女数这样的问题，如果只报告存活子女数，而没有报告平均生育子女数，有人就可能建议将平均生育子女数视为存活子女数。但是，这样一来，遇上所有这类情况时，死亡子女数都被估算为零。这会使所调查的子女死亡数也可能被估算为零，从这些数据中估算出的死亡率水平的客观性将受到损害。

121. 一致性检验将不同信息项目的值进行对照，确认前后矛盾的地方。例如显示一名14岁女子曾生育8名子女的民事登记、普查或调查记录无疑存有讹误，当然单从这些信息无法判定究竟是年龄有误，还是平均生育子女数有误。

122. 一般认为应当解决前后矛盾的问题，即

便这样做需要借助虚拟值。不过，在这种情况下使用虚拟值时，务必要确定，旧矛盾的解决不能以增加新矛盾为代价。

#### 4. 制表

123. 制表或者编制是指从构成数据集的记录中导出具有特定组合特征的人员或事件数目。“制表”一词与普查和调查连用，“编制”一词与“民事登记”连用。为了避免不必要的重复，本文使用“制表”一词。

124. 从数据集中产生的表格，应当准确地加以说明，明确指出制表的域（制表中出生、死亡或人员的分栏）、维（例如曾生子女的年龄或数目）、维的必要截尾点以及漏测值的处理。截尾点的选取通常应当确保所有案例中仅有很少一部分（例如，四分之一以下）归入未定后限的年龄组中。如出现漏测值，应在表格中单独占据相应的位置，决不能将其与其他数值合并。

125. 制表说明应当清晰详尽，不仅要让专题专家和数据处理人员明白易懂，同时使数据处理人员对表格的内容无需再费周折。

#### 5. 其他工作

126. 一旦完成计算机编辑后，可以开展其他各种数据处理工作：

- 制作印刷出版物的正稿或者制作用于数字出版的不同计算机文件可能需要的特殊处理
- 第五章第B节说明的亲生子女法可能需要对人员和户口记录进行特殊处理，使住户中的母亲和子女相匹配
- 户口调查中抽样误差的计算可能涉及密集型计算方法，因此成为数据处理的一部分
- 在人口普查中，通常应当创建一份或多份住户的样本，以备将来之用。通过这种方法，可以利用样本，而不是在全部点核记录的基础上进行补充制表和其他操作。这是处理全部普查记录之外的又一种可行方案，也可以是处理全部普查记录的前奏

## C. 存档

127. 档案库是文件和数据的储存室。存档的目的现在和过去并无二致，但是信息技术的发展已经极大地增加了存档的范围，改变了存档的手段和媒介，并且引发了有关档案安全性的新问题。

128. 以数字形式储存和处理信息的成本迅速下降，从而使我们有以较低的代价储存在任何数据收集过程中产生的大多数或者全部资料，其中包括规划资料、业务文件、调查问卷、控制表、数据集、最终结果和评估资料。

129. 现在，几乎所有这些资料都是运用计算机产生的，因而最初表现为数字形式。高效的存档要求组织好数据制作的工作，确保对数字文档进行妥善的命名、索引、加密和保存，在数据制作完成后，将文档转移到相应的数字档案库中。

### 1. 功能

130. 作为正式的参引资料，档案十分重要。数据制作机构应当贮备所有公开发布信息的权威副本。法律上可能会需要民事登记记录（有时是普查记录）的档案。

131. 档案提供了“机构记忆”，——一种对于机构过去经历的可靠的系统化记录。在规划和评估时可能需要查询这些记录。

132. 档案储存了供将来使用的数据。官方的数据收集活动几乎总是为了多种目的。其典型情况是，在收集数据时提供的信息远远多于使用或公布的信息。存档数据常常会发挥当初收集数据时未曾料想到的用途。考虑到数据收集的成本高昂，尽可能扩大所收集数据的用途十分重要。

133. 档案为对付未来信息需要的不确定性提供了手段。考虑到制作和维护数字档案的成本相对较低，在很多情况下适宜将未来可能会有用的信息进行归档。

### 2. 内容

134. 所有公开发布的信息都应当进行归档。普查、调查和民事登记中产生的所有数据集组成的各个记录应当连同处理这些记录所需的技术文献一

同归档，例如代码簿和文件格式说明。

135. 在任何数据收集活动过程中产生的许多计划、行动和评估文件也应当归档。相关项目包括，例如，各种行动计划、培训手册、计算机编辑规范和记录、制表计划和评估报告。鉴于数字媒介的储存容量基本上是无限制的，限制储存信息量多少的首要因素，是该组织是否有能力在编制文件时将文件数字化，并为文件编制索引，便于今后检索。

136. 许多实地工作的操作包括制作抽样地区的“示意地图”，如果是人口普查，则包括全国所有人口覆盖地区。理想的做法是将这些示意地图存档，然而，这并非总是可行。不过，现在人们越来越多地利用地理信息系统（GIS）来制作所使用的地图。在这种情况下，这些地图以数码形式出现，并应与其他材料一并存档。

### 3. 安全和维护

137. 档案安全包括防止丢失、讹误和越权访问。信息技术的日新月异的变化，带来了新的丢失风险，即将档案储存在行将老化的计算机媒体上所产生的风险。计算机档案必须采用相应的程序预防这种风险，定期将存档文件“更新”到新的媒体上。技术变化的速度如此迅速，计算机媒体可能在五年内就需要更新。

138. 与传统的纸张文书一样，储存数字档案的实际媒体也需要防止破坏、丢失和越权读取。由于数字信息相对便于复印和运输，它可以在不同的地点保存所有的存档材料副本，从而避免人为或自然灾害造成的损失风险。因此，在增加安全性、降低费用和极大地方便读取方面，数字档案颇具潜力。

139. 同时，数字媒体也面临着传统纸面信息不曾有过的安全隐患。由于数字信息可以轻易地更改，因而就出现了故意篡改、无意中破坏或者无意中清除造成信息完全丢失的风险，而这些风险是传统材料不曾出现的。虽然现在凭借已有的措施就能够将这些风险降低到微乎其微的程度，但是了解并实施这些措施非常重要。

### 三、 评价、估计和传播

#### 导 言

140. 评价和估计不同于上一章所讨论的操作，通常包括两个或更多的数据来源。下一节将阐述数据集评价的基本概念，包括覆盖率和内容误差以及数据质量和估计准确性的差别。

141. B节泛泛地讨论了生育率和死亡率水平及趋势估计，特别强调了从多种渠道推导和比较估计值的重要性。评价和估计密不可分，评估从一个或多个数据集中产生的估计值准确性，也是评价数据集质量的方法之一。

142. C节讨论如何以适合用户需要的形式来传播数据。过去，这主要是应当提供什么信息的问题——什么样的表格、什么样的导出统计数据，什么样的辅助文本信息，信息技术的迅猛发展同样促使人们关注信息提供的媒体和格式，尤其是计算机媒体和格式。

#### A. 评价

143. 数据制作组织应当评价它们制作的数据，并采取相应的措施，将评价结果通报给用户。如果不了解评价情况，用户可能会从数据中得出错误的结论，做出不当的决策，并根据这些决策采取错误的行动。

144. 评价对于数据收集组织的内部运作同样十分重要。这为保持和提高数据质量、根据环境的变化调整数据收集活动提供了依据。

##### 1. 活动的质量

145. 一次数据收集活动的各方面都可以加以评价，包括最初的规划和用户咨询；清单和表格的内容和格式；实地工作人员的招募、选拔和培训；表格填写完毕后的传递和控制；数据输入；人工和计算机编辑；制表计划和表格的充分性；存档和传播；以及整体规划和管理。质量控制活动记录可以

提供用于评价的许多信息。在人口普查中，普查行政报告常常对活动评价问题进行讨论。

146. 《人口和住房普查的普查管理手册》（联合国，2000年a）讨论了人口普查活动的质量评估问题。《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）讨论了民事登记和生命统计活动的质量评估问题。

##### 2. 关联性和适时性

147. 用于特定目的的数据和统计资料的有效性取决于其关联性和适时性。产生的信息要与用户的需要相关，这一点的重要性毋庸置疑，因为这是最初收集信息的原则。不过，可能有人会认为，如果在收集数据之前进行规划并与用户协商，事后就无需再开展关联性评估了。实际上，鉴于许多数据收集活动的规模、活动花费的时间以及实施活动所需要的行政系统规模，适当地坚持以用户需要为中心绝不是一件无足轻重的小事。

148. 数据的适时性十分重要，因为随着时间的推移，数据所反映的时期成为了过去，所以它们有可能会因此渐渐失去价值。在某种程度上，具备适时性就是对数据收集活动进行规划以便及时提供数据的问题。一旦启动了数据收集活动，适时性的问题在很大程度上就是确保从数据输入到最后的传播等一系列数据处理工作均遵照日程进行。可以迅速地公布初步数据，而在晚些时候发布最终的数据，通过这种方式促进适时性。

##### 3. 数据质量

149. 实地调查工作和数据处理会产生一个或多个数据集，每个数据集中的记录都包含有关构成某些统计总体的实体的信息。例如，民事登记系统每年将会产生分别代表前一年出生和死亡情况的两个数据集。人口普查中产生的数据集则是人口点查



时个人和住户的记录。户口调查产生的数据集包含调查访问时住户样本的人口和住户。下节讨论数据集质量评估时用到的基本概念。

#### (a) 覆盖质量

150. 数据集的覆盖质量指的是数据集包含的记录与这些记录应当反映的实体之间的对应关系。数据集可能会忽略一些本应反映的实体记录，而涵盖一些不应当列入的记录。如果一个数据集包含同一实体的多项纪录、未列入总体的实体记录或者不反映任何实体的记录，这就属于不当计入。

#### (b) 内容质量

151. 数据集的内容质量指的是数据集中各项记录所包含的信息的质量，由漏缺值的发生率、内容不一致的发生率以及记录包含数值的准确性来反映。

#### (c) 记录配对研究

152. 数据集评估最通用的工具是记录配对研究，将有待评估的数据集记录，原始数据集，与代表同一统计总体的次级数据集进行配对。配对的过程可能非常复杂。Marks、Seltzer和Krotki (1974) 对记录配对问题进行了颇有价值的综述。

153. 进行记录配对有两个目的。首先是将原始数据集和次级数据集中的记录分成两组，配对记录和非配对记录。配对记录是指任一数据集中的记录与另一数据集的记录所代表的实体相对应的记录。非配对记录是指任一数据集中的记录所对应的实体在另一数据集记录中未有反映。

154. 如配对程序无误，原始数据集中的非配对记录反映出该数据集中有不当计入，或者次级数据集出现遗漏。次级数据集中的非配对记录反映出原始数据集有遗漏或者次级数据集中有不当计入。因此，数据配对提供了原始数据的覆盖质量信息。根据次级数据集的不同性质，数据配对可能还会提供次级数据集的覆盖质量信息。

155. 记录配对研究所提供的覆盖质量信息价值降低，以至于两个数据集之间存在反应相关性偏向——原始数据集包含的实体更有可能在次级数据集中出现。如果发展到极端的完全相关，则原始数

据集包含的实体无一不在次级数据集中出现。严格的独立性可能很难实现，但是要想使记录配对研究得出关于覆盖质量的有用信息，必须具备些许独立性。

156. 数据配对的第二个目的是找出原始数据集和次级数据集中代表同一实体的一对配对记录。对照这两项记录能够得出有关两个数据集的内容质量的信息。

#### (d) 普查中的查点后调查

157. 进行查点后调查是为了评估普查数据的质量。查点后调查与通常的户口调查相比，存在两个技术性难题。首先，通过查点后调查，抽样的住户和人口有可能与普查中的住户和人口相匹配。这需要仔细慎重地对待调查问卷的内容，同时要在普查后不久即开展调查。第二个技术难题是，查点后调查的设计应当保证在调查中的计入机会不受普查中的计入机会的限制。《人口和住房普查的原则和建议，第1次修订本》（联合国，1998年a）讨论了普查中的查点后调查。

#### (e) 民事登记的配对研究

158. 在民事登记中，次级数据集可能源自专门为此设计的回溯性调查，源自人口普查中包含的特殊问题，或者现有的渠道，例如报章中关于出生和死亡的报道。报纸中的报道显示了反应相关性偏向的问题，因为在很多情况下报纸报道的出生和死亡与尚未报道的出生和死亡相比，登记在册的可能性更大。第四章D节讨论了登记完整性的评估。另见Marks, Seltzer和Krotki (1974年)的论著第二章以及该章的参考书目。

#### (f) 户口调查的覆盖质量

159. 在户口调查中，覆盖质量的问题略有不同。这些调查通常旨在提供相对数字估计，例如手段和比例，而不是人口或者事件的总数。户口调查中的主要覆盖问题通常是住户回答率，即提供信息的抽样住户百分比。回答率达到百分之九十或以上通常被视为可行。在住户调查中，不回答的选择性通常比不回答的水平更重要。

#### (g) 复核调查

160. 在普查和调查中，也可以利用复核调查来评估数据集内容的质量，而同样的原则也可用于民事登记。复核选取数据集中的实体样本，通过比较复核信息与原始信息来评估对于共同项目的报告。

#### (h) 统计分析

161. 数据质量的第三种评估方法包括分析从接受评估的数据集中推导出来的统计数据。数据集的净覆盖误差是指遗漏和不当计入之间的差别。可以对照普查中点查的人数和普查基准时间人口估计值，以此来估计人口普查的净覆盖误差。估计总人口的途径之一即汇总(a) 前次普查的人口估计总数，(b) 两次普查间隔期间人口的出生和死亡差数以及(c) 两次普查间隔期间迁入和迁出人口的差数。这种普查净覆盖误差的估计方法有时也被称为人口分析法。

### 4. 估计的准确性

162. 制作数据集的目的在于估计各种人口统计数量，例如出生、死亡和人口统计以及各种比率和累加测度。这些统计数量具有真值，即根据完整和精准的数据确定的数值。实际上，数据总是有瑕疵的，只能产生或多或少有异于真值的数量估值。下文的B节概括地讨论了估计。

163. 估计误差反映了推导出估计值的数据集的覆盖和内容误差，以及任何与估计量相关的数据集“选择偏差”造成的误差。

#### (a) 选择误差

164. 如果无论由于什么环境性因素，现有的数据不能准确地指代相关的统计总体，这种情况下可能会出现选择误差。例如，在人口普查或者住户调查中收集有关近期死亡的信息（见第六章B节）时，向住户中的被调查者问及过去12个月或者近期其他时间家庭成员的死亡情况。即使没有出现任何覆盖或者内容误差，所产生的数据也将遗漏在普查基准时间或者调查访问日期前解体的家庭出现的死亡人数。由单身老年人组成的住户就是这类住户的典型例子。

165. 选择偏差与覆盖误差的共同特点是：所研究的统计总体中的实体与代表该总体的数据集记录之间不完全对应。然而，选择偏差来源于数据收集过程设计的内在缺陷，即使过程执行得毫无纰漏，数据集的这种缺陷依然存在。与此不同的是，数据集的覆盖质量反映了数据收集行动的执行缺陷。第五章和第六章将讨论各种生育率和死亡率估计中存在的选择偏差。

#### (b) 评价估计值的准确性

166. 评价估计值准确性的主要技巧是比较对相同数量的不同估计值。出于若干原因，可以据此对误差进行评估。不同的估计可能出现不同的误差。即使两次估计的误差趋势相同，例如民事登记和户口调查的婴儿死亡漏报，然而误差的量值也不太可能相同。通常可以获取以往的关于误差的信息。例如，民事登记系统常常会忽略大量的出生和/或死亡人数，而人口普查很少过度查点目标人口。如果获取的是对不同数量的估计，例如年度序列的时间序列，或者所有地理分区的估计，通过对照常常可以揭示一定的模式并由此得出有关误差的结论。

167. 不同的估计也可能具有互补性的误差模式。例如，人口普查和调查中关于近期出生人口的问题（第五章E节）可能在很大程度上低估了真正的出生人数，但各个年龄段的低估程度趋于相同。因此，上报事件的年龄分布近乎正确。相反，在曾生育子女数的报告问题上，年纪较轻的妇女的报告很可能比较完整，而年岁较长的妇女则可能会漏报。如P/F比率法所示，对比这两种类型的信息，可以对两种误差做出推断（联合国，1983年）。

168. 评估生育率和死亡率的估计误差时，也可以将它们所隐含的人口增长与两次普查间隔期间的人口变化相比较，并相应地考虑人口迁移情况。最简单的办法是，将人口增长率（总出生率减去总死亡率得到的估计值）与通过后来的人口普查计算出的增长率相比较。需要时可以根据人口迁移因素和两次普查的查点完整性差别进行调整。一种更有效的方法是利用人口预测，评估前后两次普查的人口年龄分布是否与两次普查间隔期间年龄别出生率和死亡率估计一致。

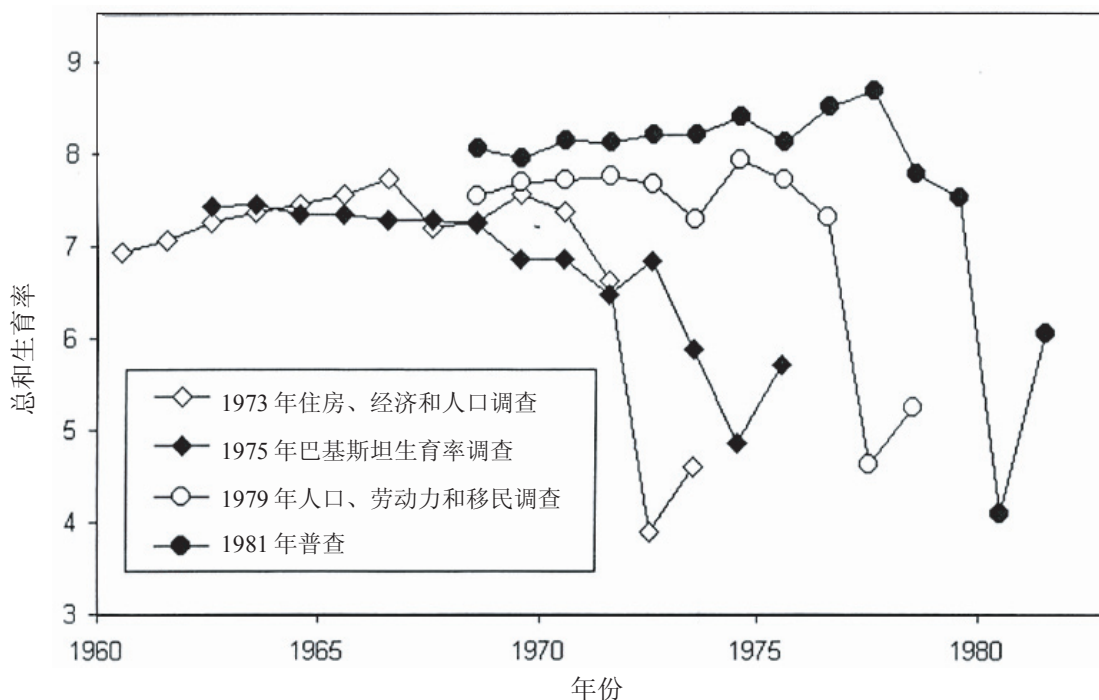
(c) 回溯性时间序列的对比

169. 如果人口普查和调查用作生育率（第五章）和死亡率（第六章）数据的来源，产生回溯性时间序列的方法就显得尤为重要，部分原因在于借助这些方法可以很容易地与所有已知的估计值进行对比，还有一部分原因是在回溯性序列中误差的模式常常为评估提供了方便。

170. Retherford等人（1987年）的分析阐述

了如何利用回溯性时间序列的比较来评价对总和生育率（TFR）的估计。在巴基斯坦的四次数据收集行动（1973年住房、经济 and 人口（HED）调查，1975年巴基斯坦生育率调查（PFS），1979年人口、劳动力和移民（PLM）调查以及1981年普查）前的15年，采用亲生子女法（第五章B节）来估计每年的总和生育率。图1绘制了这些估计值。每次数据收集行动前五年以上估计值用直线连接，以便更加清楚地反映估计数分布。

图1. 巴基斯坦总和生育率估计值比较



171. 第一次调查即1973年住房、经济和人口调查的估计显示：生育率在1960年代末开始下降，1971年以后下降速度陡然加快。其中1971年至1972年的下降速度之快令人费解。随后在1972年至1973年期间，生育率上升，这说明图上显示的下降趋势至少有部分可能由于估计误差所致。

172. 另外三次数据收集行动的估计进一步加深了这种怀疑。1975年巴基斯坦生育率调查估计显示：1971年到1972年期间，生育率并未下降，然而从1972年至1974年却骤然下降，随后在1974年到1975年之间生育率略有回升。1979年人口、劳动力和移民调查估计显示：1975年之前生育率没有显著的下降，然而1977年到1978年间却不可思议地骤

降，继而从1978年到1979年间又略有回升。1981年人口普查估计显示：1977年以前生育率大体呈上升趋势，1979年到1980年间急剧下降，而1980年到1981年间又有较大的回升。

173. 图1显示，四次数据收集行动显示的生育率下降趋势，单独来看都具有欺骗性。表面的下降源于某种误差模式大大降低了调查前第二到第三年的估计值。对四次调查的回溯性估计值进行比较，很明显在1980年以前生育率并未下降。事实上，据估计在1960年代到1970年代期间，生育率可能呈稳步上升之势。这大概是由于寡居妇女和母乳喂养减少所致。

174. 如果存在年龄虚报的问题，图1的模式

就会频频出现。Retherford和Alam (1985年)注意到,尼泊尔和印度尼西亚也出现了同样的模式。如果根据普查和调查数据反推以前的生育率,按一岁分组的儿童报告人数的波动转换为普查或调查前年份出生人数和生育率水平的波动。对幼龄儿童的查点不完整,也可能造成这种情况。

175. 对于根据平均生育数和存活子女数据产生的婴儿和儿童死亡率估计值(第五章A节),以及根据尚存父母和/或兄弟姐妹数据产生的成人死亡率估计值(第五章C和D节),均可用同样的一般方法进行质量评估。有关婴儿和儿童死亡率估计的概述,见Feeney的论著(1991年)。

### 5. 误差的指示值

176. 误差的指示值是任何说明或者暗示数据误差的统计数字或者观测值。误差的指示值常常提供次要但很有说服力的证据。

177. 以一岁分组的人口总数的分布为例。人口分布的调查发现,年龄尾数为“0”或者“5”的人数远远超过了相近年龄的人数,特别是在年龄较长的情况下。这有力地说明许多人将年龄“凑整”为相邻的5的倍数(并不一定是最相近的),这种不准确的年龄上报方式被称为“年龄堆积”。

178. 然而,这种年龄堆积的指示值并不会提供任何关于以一岁分组的各年龄组实际人数的现成估计,实际人数的估计取决于年龄点查的差别完整性以及年龄虚报的情况。

### 6. 抽样误差

179. 上述有关统计数据准确性的讨论既适用于民事登记也适用于人口普查数据——二者提供(原则上)完整的出生、死亡和人口计数,还适用于户口调查数据。然而在户口调查数据中,还需要考虑由于采用抽样而招致的额外误差。按照抽样的术语,上述讨论的误差都属于非抽样误差。

180. 原则上人们可通过样本设计来控制抽样误差并根据抽样理论来估计误差,实际操作在很大程度上也是如此。抽样误差的计算是评价和分析户口调查数据时不可或缺的辅助工具。然而,必须牢记,抽样误差是除前文讨论的误差之外的又一种类

型的误差。利用户口调查数据时,对两种误差一定要同时加以评估。

## B. 估计

181. 如果数据完整而且准确,在这种理想情况下,出生、死亡和人口的计数将通过制表产生,比率和累加测度则通过不同的定义性公式计算产生。在实际情况下,数据总是不够完整、不够准确,这时通常需要采用各种权宜手段来产生所需要的人口统计数据。

182. 在这里,估计指某些根据现有数据无法准确计算的人口统计量的数值推测手段。结果就产生了对数量的估计值,即预计或多或少会偏离真实数量的数值。即使能够获取完整而准确的民事登记和人口普查数据,比率的计算也需要对分母进行估计。

### 1. 民事登记数据

183. 如果民事登记数据完整而且准确,对其提供的出生和死亡计数无需进行调整或者估计,即可运用于比率和累加测度的计算。不过,在许多国家,尽管民事登记是弥足珍贵的数据来源,然而由于其完整性和准确性均不高,不经调整无法使用。在这种情况下,可根据登记在册的出生和死亡人数,以及从人口普查和户口调查中获取的辅助性信息,来估计出生和死亡的真实人数。随后再利用这些估计的真实值来计算比率和累加测度。

184. 民事登记可能会提供完整的出生和死亡数据,然而,它所提供的有关出生、死亡特征的信息却可能存在缺陷。例如,死者死亡的年龄、死因、或者出生时母亲年龄或已生育子女数。实际操作中,覆盖质量和内容质量之间势必存在对应关系。尤其是许多出生和死亡登记不全的国家,关于死亡年龄和死因的报告常常存在误差。

185. 第四章D节讨论了出生和死亡民事登记报告的完整性评估。

186. 如果民事登记数据完整而且准确,唯一需要进行估计的就是用于计算比率的分子。如果分子的信息来源于人口普查(户口调查也许作为辅助信息来源),应当估计所需各年份或者其他时期、

各人口亚群（例如年龄组）的比率的分子。第四章B节第3小节讨论了分母的估计方法。

## 2. 普查和调查数据

187. 利用人口普查和户口调查产生生育率和死亡率数据大概有三种方式。第一种方法是利用有关出生和死亡的回溯性问题。这种方法试图通过普查或者调查获取与民事登记所提供的相类似的信息。最简单的例子就是向育龄妇女提问，询问她在普查基准时间或者调查访问日期前12个月是否生育过子女（第五章D节）。了解生育史则比这更加细致一些，应询问妇女她们所有子女的出生日期以及其他详情。

188. 第二种方法仅提供有关生育率的数据。这种方法利用每次人口普查和户口调查收集到的年龄信息，估计普查或者调查之前年份的出生人数。这种方法依据一个简单的道理：在普查时零岁（以周岁计）的人口必然出生于普查前一年。以此类推，一岁的人口则必然出生于普查基准时间前两年。估计出婴儿和儿童死亡率水平之后，就可以针对普查前夭折的已出生儿童人数进行调整。这是存活倒推法、亲生子女法以及生育史重建法的依据。第五章A、B和C节分别对这些方法予以阐述。

189. 第三种方法是针对生育率和死亡率的情况提问，但无意了解关于出生和死亡时间的具体信息。这种最古老、应用范围最广的方法所依据的问题是：截至普查基准时间或者调查访问之际，妇女既往生育子女的总数以及尚存的子女人数。很明显，尚存子女的比例说明了死亡率水平，不过这种信息如何用于估计死亡率水平还不明确。这种方法的其他例子是有关尚存的父母（“孤儿”问题）以及兄弟姐妹存活的问题。这三种方法分别在第六章A、D和E节予以讨论。

190. 过去的半个世纪里，关于如何根据普查和调查数据估计生育率和死亡率水平和趋势，发展了许多技术。《手册十：人口估计间接技术》（联合国，1983年）以及即将出版的有关成人死亡率估计手册对估计方法进行了说明。有关这个问题的其他参考资料分别在第六章和第七章列出。

## 3. 根据多次估计值产生的水平和趋势

191. 不同数据采用的不同估计方法通常会得出对同一数量的不同估计值，例如总和生育率（对照上文A节第3小节）。在某些情况下，对同一数量的多次估计几乎是相同的，而在另外一些情况下，却有较大的出入。

192. 如果出入较大，有两种可能性。可能某种估计或某个系列的估计明显优于其他估计，如果属于这种情况，这种估计或者系列就会被甄选为最终估计。另外一些情况下，原始估计或者系列中没有一种明显地出类拔萃。在这种情况下，较好的方法是不从这些估计中挑选，而是汇总所有估计中包含的信息。

193. 如果涉及到某一数量的多次估计值，也许可以通过计算平均值、加权平均值或者中位数获取最终的估计值。如果涉及到不同时间序列的估计，最好的办法常常是将一条直线或者曲线与所有现成的序列拟合。有时候，如果现有的估计值准确度不高，最合理的方法就是用一条拟合直线反映所估计数量的水平和趋势（例如，总和生育率、出生预期寿命、婴儿死亡率）。如果估计值的准确度较高，则适于运用曲线拟合。

## C. 传播

194. 将收集到的信息以符合用户需要的形式提供给潜在的用户，这时数据制作工作才算完成（联合国，1998年a）。传播的目的就是向可能会用到信息的人员或者机构提供信息。

195. 准确地说来，传播是系统性、持续性地找出数据的潜在用户，告知他们可以读取的信息并从他们那里获取有关信息使用和信息需求的反馈意见。数据的制作者和使用者之间持续性的沟通，是确保所生成数据的价值和利用的最有效途径。

### 1. 印刷出版物

196. 全国性统计机关或者负责该项工作的其他实体能够利用的印刷出版物属于传统的传播形式。除此之外，还有各种打印报告，例如工作文件或者适当标注和编目的计算机输出资料，这些报告针对少量受众，对于他们来说，传统的出版物缺乏成本效率。

## 2. 数据集

197. 包含出生、死亡和人口记录的数据集对于用户而言十分宝贵，因为这样用户就不会拘泥于根据公布的图表做出的分析。这是非常明显的优势。因为从大型数据集中可以制定的图表数目十分巨大，而数据制作机构只会生成很少一部分表格。

198. 数据集的传播对于数据制作机构十分重要，因为这是传播数量巨大的信息时最具有成本效率的方式。如果用户自行生成表格，就减少了表格制作和传播的费用。

199. 人口普查和民事登记数据集的传播引发了个人保密性的问题。普查和民事登记数据的保密性通常受法律的制约。无论是否存在法律规定，保密性对于数据制作机构都具有根本性的意义，因为这影响了数据制作机构从答卷人和申报人那里获取准确信息的能力。因此，在公布数据集之前，必须对数据集进行相应的处理，隐蔽个人的身份。有时候，这一过程被称为匿名化。

200. 对于人口普查和民事登记数据集，需要同时运用抽样和删除具体地理位置信息的方法以确保保密。对于户口调查，则只需要采用后一种方法。保密性的基本标准是使从数据集中辨别个人身份的可能性降低到几乎为零。

201. 最后的数据集常常称为公用样本。普查数据的公用样本常常是计算机抽取的住户样本。其中可能包括或者不包括“集体”户，例如监狱犯人以及其他机构的居住人员。民事登记数据的公用样本是已登记的出生和死亡样本。

202. 由数据集的传播引发的第二个问题是用户实现潜在价值的的能力。过去，对于人口普查或者民事登记系统生成的大量数据，用户很少具备储存和处理的设备能力。随着信息技术的飞速发展，越来越多的用户具备了处理大量数据集的能力。这导致人们对数据集的需求增加，在可以预见的未来这一势头估计仍将持续。

203. 目前，大国的人口和民事登记数据集仍然超出了大多数用户的利用能力。未来十年，对用户能力的限制可能会消失，然而保密性问题仍将存在。尽管有可能对人口普查或者民事登记数据集中

的每一项记录进行匿名处理，但是这些数据集的完整覆盖价值大多取决于它们提供的详细地理信息。这类数据集的主要效用可能是小型人口亚群的研究，例如少数民族。

## 3. 数字出版物

204. 随着个人电脑的兴起，越来越多的用户利用计算机处理各类数据和信息。如果数据仅仅以印刷品的形式出现，它们必需转化为数字形式方能使用。这是一项耗时又费钱的工作，而且容易出错。如果数据的数量非常巨大，成本常常是惊人的。另外，这本身就是一种浪费，因为现在几乎所有的数据都通过计算机生成，因而从一开始就是数字形式。

205. 数字出版物指以适于计算机处理的形式传播的信息。由于数字出版物的出现，用户不必再将印刷资料转化为数字形式。通过这种方式，数字出版物节省了用户的时间、精力和成本，提高了数据的利用。

206. 数字出版物可能需要分发实际的计算机媒体，包括含有所传播信息的计算机磁盘、磁带或者CD-ROM。在这种情况下，传播中需要动用与印刷出版物相同的运输网络。数字出版物也可以通过计算机网络发送，例如因特网，这种情况不涉及实际媒体的分发。信息将直接从一台计算机的内部存储器传送到另一台计算机的内部存储器。

207. 当信息通过计算机网络传输时，发送可以在线也可以离线进行。离线发送意味着负责机构的某指定人员受理用户的数据申请。数据申请可以通过计算机网络发送，但是处理申请需要人力干预。在线发送也同样涉及用户申请和应申请而传输的信息，但是计算机服务器可在没有人力干预的情况下处理申请。

208. 计算机网络传播为实现在线制表提供了可能。这时，用户通过网络向数据收集机构或者其他负责实体管理的服务器提交制表申请。通过程序设置，服务器可回复这些申请：首先核查制表申请是否符合用户保密性的制表原则，如果申请可以通过，则生成表格并通过计算机网络将表格发送给申请人。服务器可设置为自动实施上述操作，而无须人力干预。人口与健康调查的STATcompiler对在线制表予以了

说明,可在人口与健康调查网站上查阅:<http://www.measuredhs.com>。

209. 目前,互联网是最方便的在线发送设施(它亦可用于离线发送),不过也可以通过ftp(文件传送协议)来进行传送,或者通过计算机公告牌,即用户可以通过电话读取和检索该计算机上的数据。

#### 4. 数字格式

210. 一台计算机上生成的计算机文档是否可以用于其他计算机,这首先取决于文档的格式,其次是相关计算机的属性。有些文档格式几乎通用于所有的计算机,这样的格式被称为独立平台。其他格式只能在应用特殊软件后方能使用,因此限于特定的计算机操作系统。

211. 有些格式的文件公开,供所有人免费使用。其他格式则是专有的,文件不公开,人员或者机构在获取使用许可证后方可使用。这类格式通常需要特殊的计算机程序。

212. 目前美国信息交换标准代码(ASCII)格式的文本文档提供最接近国际化的格式,具有平台独立性和非专有性。文本文档不同于二进制文档,后者特别针对特殊的计算机应用软件(和操作系统),在其他应用程序中常常无法读取。Unicode全球字符标准正在扩充ASCII格式,该标准支持近50 000个不同的字符,包括世界上主要的书写文字和符号系统(Unicode 协会,2000年;又见<http://www.unicode.org>)。

#### 5. 数字复印

213. 印刷出版物的历史十分悠久,数字出版物在这样的背景下迅速地涌现,因而将数字信息储存与传统的面向页面的表现形式相结合的计算机文档格式层出不穷。在计算机屏幕上出现的页面与打印的外观相同。这种类型的文档可以称为“数字复印”或者“数字传真”。这类格式中应用最广泛的是Adobe公司的可移植文档格式(PDF)。

214. 数字复印格式的优势在于它使得出版人和用户能够利用数字媒体的某些优势,同时又遵循印刷出版物为人们所熟知的规范。有了数字复印格式,出版人能够将文档传送给用户,而用户打印的

文档与它们原本可能会接收到的印刷出版物完全相同。考虑到所需的技术基础设施,所发布信息的储存、复印和运输或传输成本大大降低。确实,过去信息的售价常常需要足以抵补印刷、储存和传播的成本,而现在则可能会免费发放。

215. 印刷出版物的面向页面特性,从一方面来说,是数字复印格式的巨大优势,而从另一方面来说,这恰恰也是一个不容忽视的劣势,特别是在当出版物的内容主要由表格数据构成的情况下。尤其是普查制表,大小悬殊较大,小型的表格只需要页面的一部分,而大型的表格可能会占用数十页甚至数百页。印刷媒体要求页面大小固定,因此大型表格必须按照页面大小进行分割。如果表格主要供用户进行肉眼核查,则劣势可能还不很明显。然而,如果表格需要由计算机进行处理(特别是大型的表格,几乎总是会进行这样的操作),整个表格必须从分割的页面中重新组合。

216. 其他数字文档格式的强大优势在于它们不要求含有表格的“页面”大小统一。每张表格,无论是大是小,都可以按照原本的尺寸安置。这极大地简化了计算机对表格的处理。如果是用肉眼核查,与计算机屏幕不相称的表格可能需要“滚动”。因此,在这种情况下,数字复印格式的面向页面特性不仅不是优势,反而成为了负累。

#### 6. 标准通用置标语言(SGML)与可扩展标记语言(XML)

217. C节第6小节讨论的工具旨在解决以数字形式向用户传播数据和相关文档时遇到的几个基本问题。

- 提供在不同计算机操作系统下运行的各种计算机软件应用程序都可以方便地读取的信息
- 提供“自编文件”数据——共存于一个方便的数字包中的数据和文档
- 提供恰当形式的信息,使用户通过计算机预处理信息的能力最大化,例如,尽可能地减少人工搜寻大量文档所需的工作量

这些工具也可以用于数据制作机构的信息管理。例如,它们可用于下述用途:

- 实现文档管理系统——该系统用于储存文件、跟踪文件内容、控制文件读取，同时方便迅速查找文件所含的信息
- 储存、公布和执行各种标准，例如数据文件制作标准，以及标准编码方案，例如地理区域和死因的编码方案
- 实现以标准格式储存信息的“多元化利用”系统，从该系统中可以方便地生成其他多种格式，包括印刷出版物的面向页面的格式，互联网上发送的超文本标记语言（HTML）文件，以及試算表文件（针对希望以这种形式处理信息的用户）

218. 在描述SGML和XML之前，应当首先说明推动这些工具发展的综合动力。如以上的C节第4小节所示，在大多数计算机运行系统中，多数软件应用程序都可以读取ASCII格式的文本文件。文本文件也可能包括文本注解数据和数据本身。这些特征构成了非常有利的优势。

219. 文本文件也存在不利因素（二进制制文件也是如此）。计算机所谓的“读出”文本文件，只不过是识别文件中包含的字符。计算机无法确定文件内容的含义，而只有人类才能真正地从人的意义上读出文件。

220. 为了阐明这一差别的重要性，假设某个用户希望了解某国的哪一次人口普查提出了平均生育数的问题（见第五章C节）。另外假设，在这个例子中，互联网上以普查问卷的传真件形式提供了这一信息。

221. 为了获取所需要的信息，用户必需找出互联网上每次普查的调查问卷，并阅读问卷，查核其中是否包括有关平均生育数的问题。熟悉印刷媒体的用户可能无法想象还有任何其他方案，然而，习惯于使用互联网搜索引擎的用户则会寄希望于计算机来进行这项操作。

222. 那么，假设用户需要的不是一国，而是多国的信息，甚或是全世界将近200个国家的信息。一方面，这需要投入无比巨大的人力，而另一方面，这正是计算机应当帮助用户开展的一项例行的文书工作。

223. 想要让计算机向用户提供这项服务，必须创建代表普查、调查或者民事登记表的文件，计算机可以向人类读者一样识别文件中的信息。目前常规的操作办法是用代表不同信息的“标记”来“标注”文本。例如，下面是一段表述普查问卷的标记文本：

## 方框 2. 可扩展标记语言（XML）图示

```

<普查_调查问卷>
  <国别> 南非 </国别>
  <年份> 2001 </年份>
  <时间> 10月 9-10 日午夜 </时间>
  .
  .
  .
  <问题>
    <编号> P-03 </编号>
    <名称> 性别 </名称>
    <措辞> （该人员）为男性或者女性？ </措辞>
  </问题>
  .
  .
  .
</普查_调查问卷>

```



角括号中的文字，例如〈普查\_调查问卷〉，就是标记。所有其他文字是内容。不同的内容被合围在指示内容含义的标记中，标记方便了文件的自动处理。

224. 在这个例子中，标记的语法和含义在很大程度上是不言而喻的。标记包含了说明内容含义的名称，而所有的内容都包含在起始标记和结束标记中。例如在标记〈问题〉和〈/问题〉之间所包含的内容指的是调查问卷上的问题，标记还可能包含识别信息和对查点员的说明。标记可能是嵌套式的（包含其他标记）。例如，问题包含三个标记成分：调查问卷上问题的数目、问题的名称以及问题的措辞。

225. 如果调查问卷包含在文本文档、文字处理文档或者传真图像中，计算机程序无法确定问卷中是否包含特定的问题。如果含有调查问卷的文本文档经过了适当的标记，计算机就能够完成这项操作以及其他许多有关问卷内容的信息处理任务。

226. SGML代表**标准通用置标语言**——一种国际标准，ISO 8879。可以利用该标准定义多种不同用途的置标语言（Goldfarb, 1990年）。XML代表**可扩展标记语言**，是从SGML中推导出来的一种简单而灵活的文本文档格式。XML的发展与互联网的应用程序相关。互联网联盟的网站（<http://www.w3.org>）提供有关XML、相关说明及其用途的详细信息。

## 四、 作为生育率和死亡率数据来源的民事登记

### 导 言

227. 民事登记的首要目的是从法律上证明出生、死亡和其他生命事件；提供统计信息只是它的一个次要功能。不过，当民事登记完整地覆盖了人口的出生和死亡等事件时，它就成为关于生育率和死亡率的基本数据以及关于死亡原因的数据的优先来源。

228. 本章描述了如何将民事登记用作生育率和死亡率的数据来源。A节就民事登记系统的运作做了简要说明，主要强调了民事登记的数据收集和处理方面而不是法律方面。B节阐释了如何使用民事登记数据计算生育率和死亡率。

229. C节提出了登记的完整性的概念。D节论述了当生育率和死亡率登记不完整时如何使用民事登记数据。不完整的民事登记数据常常是关于生育率和死亡率信息的一个重要来源，但其使用需要特殊的方法。尽管如此，使用不完整的民事登记数据除了能够提供关于生育率和死亡率的信息之外，还能够推动民事登记的不断发展。

230. 本章的最后一节，E节，探讨了如何将人口登记与民事登记系统结合起来，用作生育率和死亡率的一个数据来源。

231. 关于民事登记系统的详细信息见《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）以及五卷的《民事登记和生命统计系统手册》（联合国，1998年b，1998年c，1998年d，1998年e和1998年f）。

#### A. 民事登记

232. 民事或生命登记的定义是连续、普遍、长期和强制性地记录生命事件的发生和特征，进行民事登记主要是为了制定法律规定的法律文件，其次它们也是生命统计的有用的资料来源。生命事件包括活产、死亡、死胎、结婚、离婚、收养、私生

子的合法化和承认、婚姻无效和法律判决的分居（联合国，2001年b）。

233. 民事登记有着深刻的历史根源。各类政府登记簿的存在已经有几千年的历史。某些发达国家的民事登记系统也已经使用了200多年。因此，这方面大量的经验、概念、定义和程序均已有了相当完备的发展，并已进行了大量的整理编纂。

234. 各国对民事登记做出规定的法律均有不同，但是为了促进发展有效的民事登记系统，且为了各国之间的可比性，各国应尽可能地遵守由联合国出版的关于民事登记的各项原则和建议，并适当考虑本国国情（联合国，2001年b）。

235. 对民事登记数据的统计编纂应该进行中央计划、组织和实施。在某些国家，统计编纂工作和民事登记系统业务在组织机构上可能是分开的。

#### 1. 中央民事登记机构

236. 一国关于民事登记的法律应当指定一个中央机构来主要负责民事登记。该法律应规定该中央机构的具体义务、权力和责任。在联邦制的情况下，中央机构不妨设在省或州一级，而不要设在联邦一级。

237. 一般说来，中央机构将负责规定主要登记区，并创建、运作和维护为主要登记区服务的地方登记处网络；负责建立一个汇编地方登记处为统计目的收集的信息的系统；负责对所收集信息进行安全和保密管理；负责向网络的所有组成单位提供技术指导，其中包括为登记工作人员提供培训。中央机构还将负责与其他支持民事登记系统的政府机构进行协调，例如负责卫生保健服务和统计工作的机构。关于中央机构的作用的详尽信息，见《关于生命统计系统的原则和建议》（联合国，1973年和2001年b）。

## 2. 地方登记处

238. 民事登记需要在全国建立登记处，登记处数量应当充足，地点应当合适，使得每一个住户与地方登记处都在适当距离之内。地方登记处应当配置带薪登记员，他们应当经过登记程序方面的培训，备有必要的表格和参照信息，并有充分的时间从事工作。登记处的办公时间应当合适。

## 3. 登记程序

239. 由一名依法指定的申报人和地方登记员负责登记生命事件的工作。申报人负责报告发生的事件并提供关于事件的事实情况。地方登记员负责汇编并处理根据申报人提供的信息所做的登记记录，如果适用的话，还负责核可内科医生或其他核证官所开具的事件证明。

240. 为了民事登记的目的，一国的地理区域被划分为主要和次要登记区。可在一个主要登记区内授予次要登记区权力，特别是在大医院和其他发生出生和死亡事件的卫生保健中心。一般说来，各登记区由一个地方登记处向它提供服务，但是偏远地区和人口稀少地区可由流动登记站提供服务。

241. 主要登记区应当对那里所发生的出生和死亡事件进行登记。应当对主要登记区的界限进行明确界定和划分，以免对登记任何出生或死亡事件的地区辨别不清。

242. 申报人由有关民事登记的法律具体规定。就出生事件而言，申报人可以是母亲、父亲或接生医院的行政官员。就死亡事件而言，申报人可以是最密切的尚存的亲属或死亡时在场的专责医生。为尽可能减少对由谁负责问题发生混淆，建议法律明确指定一个人作为每一个事件的申报人。另外还应当指定一名替代申报人，以备在主要申报人不能履行其责任的情况下履行这一责任。

243. 建议民事登记法规定在事件发生几天或几周内必须按规定时间对事件进行登记。法律可以规定对超过这一时间进行登记的事件收取小额罚款。但这种制度应当允许延迟登记，采用的方式最好是劝止今后不要延迟出生和死亡事件登记，但又不致使人们不愿为已发生的出生和死亡事件进行登记。

## 4. 出生和死亡登记记录

244. 登记出生或死亡事件时获取的信息由出生登记记录或死亡登记记录界定并记录在这两项记录上。可在合法登记记录的内容与为统计目的收集的补充信息之间划分一条界线。登记记录由当地登记员在登记时作出。

245. 为产生生育率数据目的，最重要的出生登记记录项目包括如下：

- 发生日期和地点
- 登记日期和地点
- 胎儿类型（单胞胎、双胞胎等）
- 分娩时专责医务人员（医生、护士等）
- 出生地点类型（医院、家里等）
- 孩子性别
- 出生时体重
- 母亲出生日期
- 母亲常住地址
- 母亲的活产子女数
- 母亲的婚姻状况
- 母亲当前婚姻的缔结日期（如果已婚的话）
- 母亲前一次活产的发生日期（如果有的话）
- 母亲最后一次活产的发生日期（如果有的话）
- 母亲最后一次行经日期
- 父亲的出生日期

246. 关于建议列入的有关出生记录的完整项目清单，见《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）。该清单不包括最后一次活产的存活情况。但是，在婴儿和儿童死亡事件登记不完整的情况下，最后一次活产的存活情况可以提供关于婴儿和儿童死亡率的宝贵数据（Brass和Macrae，1984年和1985年）。

247. 为产生死亡率数据，关于死亡登记记录的最重要项目应当包括如下：

- 发生的日期和地点
- 登记的日期和地点
- 基本死因
- 死因证明人

- 死者的常住地址
- 死者的出生日期
- 死者的性别
- 死者的婚姻状况
- 死者（婴儿死亡）母亲的常住地址

关于建议列入的关于死亡记录的完整项目清单，见《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）。

248. 基本死因的定义是：引起直接导致死亡的一系列事件的疾病或伤害，或造成致命伤害的事故或暴力情况。有关死因的详情，见《疾病和有关保健问题国际统计分类，第十次修订本》，第2卷，《指导手册》（世界卫生组织，1993年）。

#### 5. 民间社会的了解和责任

249. 除了必要的政府基础设施以外，民事登记还需要公众普遍对民事登记有所了解、能够进行登记并给予合作。尤其是，申报人必须了解他们已经被指定从事这项工作并且知道去哪里登记他们负责登记的事件。他们还应当知道事件应当及时报告并准确了解登记时所需提供的信息，例如死者在死亡时的年龄。信息、教育和宣传（IEC）方案的目的是提高公众对民事登记的了解，并促进所有生命事件的登记，这些方案对民事登记系统的发展和维护至为重要（联合国，1998年b）。

#### 6. 制表方案

250. 在一个充分完善的民事登记系统中，每一项出生或死亡事件都在发生数周内向当地登记处进行登记。事件的统计记录随即被纳入列有所有已登记的出生事件或所有已登记的死亡事件的中央数据集。出生和死亡表可以根据这些数据集按周、月、季度或年度编制。期间越长，出生或死亡表的内容就越详尽。

##### (a) 发生日期和登记日期

251. 任何期间的最后列表一般都应当基于各个期间发生的事件，对照该期间登记的事件。为了确保随时间的可比性，各个期间的表格都应当在统一延迟的时间以后编制，例如一月份的表格应当在

二月底编制，各个历年的表格应当在历年结束后三个月编制。

252. 在急需得到尽可能当前的数据时，可以根据登记日期编制特殊表格，只要最近信息的价值不会因为导致覆盖面误差和选择偏差而变得无效。导致覆盖面误差是因为在任何期间登记的事件(a) 包括一些发生在早些期间的事件，而(b) 没有包括一些在特定期间发生但在该期间结束时没有得到登记的事件。净覆盖面误差是这两组事件数量之差。即使净误差为零，在该期间登记的事件的各种特征的分布也不会与在该期间发生的事件的相应分布相同。

##### (b) 发生地点和登记地点

253. 出生数和死亡数可按事件发生地点和与事件有关的一个或多个被调查人的常住地址进行分类。居住地址可能会随时间而变；给出的居住地址是指截至事件发生时的地址。在出生表中，被调查人是母亲。在所有死亡表中，被调查人是死者。在婴儿死亡表中，被调查人是死亡儿童的母亲。

254. 发生地点与登记地点的区分适用于国家一级制表的制表，也同样适用于国以下一级的制表。一国的公民有时在国外生育或死亡，其他国家的公民有时在正在为之制表的国家生育或死亡。民事登记系统可以采取对居住在国外的公民的出生和死亡事件进行登记。

255. 然而，区分发生地点与登记地点对编制提供国家的地理分区的数字的表格极为重要。总的说来，地理区域越小，区分就越显重要，因为发生地点与居住地点之间的距离分布主要集中在距离相对较短的地方。

##### (c) 为国以下地理区域提供列表

256. 民事登记作为生育率和死亡率数据一个来源的一个重要好处是可以为国家每一个地区提供数据。这便于为许多不同地区分类系统提供列表。

257. “地点”的具体意义取决于所采用的地区分类系统。《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）建议为农村地区、主要城市和城镇以及大小民政区提

供列表。

258. 正如第一章A节第4小节中所述，为其他地区分类系统制表对行政、计划和政策目的可能十分有用。《地理信息系统和数字化地图手册》中所举的例子包括劳动力市场区、经济发展区、学校区、交通规划区、卫生保健规划区和水资源规划区（联合国，2000年b，图三.1）。

#### (d) 年龄别

259. 出生和死亡表应当始终包括年龄别列表。提供充分的年龄组详情非常重要。对于出生事件来说，年龄指的是母亲在生育时的年龄。对死亡事件而言，年龄是指死者在死亡时的年龄。

260. 按母亲年龄分列的出生表一般将至少包括一个国家一级的按一岁分组的表格，对15岁以下和49岁以上（通常认定的育龄期）的妇女的生育率规定了适当的容许差。其他表格一般仅仅显示标准的按五岁分组的年龄组，但对不在常规育龄期之内的妇女的生育事件也同样规定了容许差。不明年龄的报告应当始终填在“不确定”类别。

261. 按年龄分列的死亡表一般应当至少包括一个国家一级的按一岁分组的表格。按五岁分组的死亡表应当始终将0-4岁年龄组分解为0岁（婴儿死亡）和1-4岁。最好把1-4岁年龄组也分解成按一岁分组的年龄组。0-4岁年龄组里按年龄分列的差别死亡率始终是极端的。如果对这个年龄组不进行分解，重要的信息就会丢失。未定后限的年龄组的起点年限要高，足以使属于这一年龄组的死亡人数可以忽略不计。建议采用的未定后限的年龄组为“100+”（100岁以上）。

#### (e) 制表的实例

262. 可能编制的表格显然受到出生和死亡统计记录中所包括的项目的限制。然而，即使内容极少，这两种表也可以分为以下四大组：

- 用于计算出生率和死亡率的表格
- 提供有关出生和死亡事件特征以反映人口的卫生健康条件信息的表格
- 用于分析民事登记系统的特征和运作情况的诊断表

#### • 杂项表

263. 表格还可以分为国家一级表格和国以下各级地区及人口亚群表。与国以下各级地区或人口亚群相比，国家一级表格普遍越来越详尽。

264. 特定的国情可能会提出特殊的要求。例如，在一个有大量非公民在那里居住和工作的国家，制表方案可以包含公民与非公民的区分。一些表格可以包括一个公民资格范围。其他表格可以对公民与非公民分别制表，可以对这两个群体编制不同的表格。

265. 用于计算出生率和死亡率的表格举例如下：

- 按出生时母亲年龄分列的出生数，用于计算年龄别出生率
- 按死者死亡时年龄及其性别分列的死亡数，用于计算年龄别死亡率和生命表
- 按母亲年龄和活产胎次分列的出生数，用于计算年龄-胎次别出生率和生育概率
- 按死者死亡时年龄、性别和死因分列的死亡数，用于计算死因别死亡率和生命表
- 按出生年份（本年度、前一年）分列的婴儿死亡数
- 按死亡事件发生的月份和以周月表示的死亡年龄分列的各历年期间的婴儿死亡数

266. 下面B节讨论了利用这些表中的前两个表计算年龄别出生率和死亡率问题。第五个表对计算婴儿死亡率很有用。最后一个表有用于计算生命表。可以通过提供以下方面的信息对这一问题作详尽阐述：如果死亡发生在生命的第一个月，则以周表示死亡时年龄；如果死亡发生在生命的第一个周，则以天数表示死亡时年龄；如果死亡发生在生命的第一天则以小时表示死亡时年龄。

267. 以下例子所叙述的表格提供了有关出生和死亡事件特征的信息，反映了人口的卫生保健条件：

- 按发生地点、出生时专责医务人员和是否住院分列的出生
- 按发生地点、出生时体重和是否住院分列的出生

- 按发生地点、妊娠期和是否住院分列的出生
- 按发生地点、是否住院和证明类型分列的死亡
- 按发生地点和基本死因分列的死亡

268. 如果发生地点分类信息更加详尽，例如按较小的民政区或较小的地区分列的话，这些表格提供的信息将对卫生保健服务的规划和行政管理极为有用。在某些情况下，用包含常住地址的选定表格补充上述清单是有用的。

269. 以下提供的是用于分析民事登记系统的特征和运作的表格实例：

- 按母亲常住地址和出生发生地点分列的出生
- 按母亲常住地址和婴儿死亡发生地点分列的婴儿死亡
- 按死者常住地址和死亡发生地点分列的死亡
- 按发生年份分列的本年登记的出生
- 按发生年份分列的本年登记的死亡

270. 前三个表提供了分析发生地点与居住地点之间差异的根据。后两个表如下文C节中所讨论的，可用于对延迟登记的分析。杂项表的例子包括如下：

- 按母亲年龄和父亲年龄分列的出生
- 按母亲平均生育数和存活子女数分列的出生
- 按前一胎次存活状况（存活、死亡）分列的第二和以上胎次的活产

271. 第一张表提供的信息有助于间接估计程序，还有用于社会研究。表二和表三提供了有关婴儿和儿童死亡率的信息。即使在出生登记极不完整的情况下，这三张表也将十分有用。

## 7. 出版计划

272. 民事登记系统的统计成分将包括定期和及时出版关于生命事件包括出生和死亡的表格数据以及适当的文献资料。文献资料包括术语的定义、各种方法和民事登记系统程序的说明以及对出生和

死亡事件登记的及时性和完整性的评价。

273. 一旦获得了国家任何实质地理区域的出生或死亡事件的实质部分，即可启动出版计划。虽然对于什么是“实质性”这个问题，本文并没有提出公认的准则，但是一个合理的解释是，在一个相当于国家首都大小的地区，登记的出生或死亡事件应当超过四分之一。

274. 在出版已知是以不完整的登记为根据的出生和死亡表时，应当同时出版相关的文献资料，以便让用户了解数据的局限性。有关登记系统在国内不同地区的运作情况的信息以及有关任何现有登记完整性的估计的信息特别重要，尽管登记目前仍然是不完整的。

275. 有时认为，除非出生和死亡登记是完整或接近完整的，否则民事登记数据毫无用处。然而，即使是很不完整的民事登记数据也可能含有宝贵的信息，出版已获得的数据将会对进一步发展民事登记起到些许推动作用。有关进一步讨论见下文E节。

## 8. 民事登记数据的质量

276. 民事登记数据的质量包括第三章A节第3小节中所讨论的覆盖面和内容范围。在国家一级，对事件按发生日期进行制表时，民事登记数据中的覆盖面质量首先是一个覆盖面完整性问题；不当计入的情况往往是极少或不存在的。应当强调的是，出生的覆盖面和死亡的覆盖面可能有很大的差异，凡在覆盖面不完整的地方，婴儿死亡登记的完整性可能大大低于非婴儿死亡登记。

277. 对事件按登记日期进行制表时，情况有很大的不同。通常对某一日历年或其他日历期间登记的出生或死亡本来就不感兴趣。按登记日期编制的表格将被认为提供的是按发生日期汇总的表格数据的近似值。因此，在某一时期发生但是在该时期结束后登记的事件为漏计，在该时期开始之前发生但是在该期间登记的事件则为不当计入。因此，事件的发生与登记之间的延迟模式的改变可能导致事件的估计数过大或过小。

278. 例如，假设在过去，任何一年发生的全部事件中有25%是在第二年登记的；但是登记制

度的改进使得这一百分比在本年减少到零。在本年登记的出生事件就包括在这一年发生的全部出生事件，加上在上一年发度生的全部事件的25%。出生事件在本年就将“重复登记”25%。

279. 在国以下一级，覆盖面误差问题在对事件按发生日期进行制表时变得更加复杂。民事登记系统中的质量控制程序应当确保准确报告登记地点。但是对常住地址可能会发生误报，这可能会给局部地区带来覆盖面误差。此外，发生地点数据有时用作居住地点数据的代用资料。这可能导致覆盖面误差的高水平，包括正数和负数。见下文B节第2小节的例子。

280. 由于民事登记的主要法律功能的缘故，为民事登记所做的统计处理工作不如为普查和调查所做的工作充分。如果出生和/或死亡的报告看来不完整，就应当努力确定统计处理工作而不是登记本身是否出错。从统计角度看，一个事件只有在以下条件下才算是得到了登记：(a) 已在当地登记处登记和(b) 这一信息已被收入用于制表和编制其他统计资料的数据集；单是登记是没有意义的。

281. 民事登记比普查和调查方法具有优势，因为信息在事件发生后不久捕获的，而且信息来自一个最可能能够提供准确信息的申报人。然而，如果在普查和调查中对出生日期或年龄的报告不准确的话，它们在民事登记数据中的报告也同样是不准确的。

#### B. 利用民事登记数据计算出生率和死亡率

282. B节论述在民事登记被认为是完整的情况下计算出生率和死亡率的问题。在实践中，这通常意味着评价研究所估计的出生和死亡登记的完整性程度很高，足以认为没有必要对没有登记的事件作出任何调整。关于不完整登记数据的使用问题在下文D节中讨论。

283. 根据民事登记和人口普查（也许还有户口调查）数据计算出生率和死亡率包括两个主要阶段。第一个阶段的目的是从民事登记中获取出生数和死亡数，并从普查和调查中获取个人的相应数字。

284. 第二个阶段是估计根据普查和调查数

据得出的出生率和死亡率的分母。人口普查通常每10年进行一次，但是所需要的分母是每年、每个季度、每个月、也许甚至是每个星期的出生数和死亡数。关于分母估计的论述见下文B节第5小节。

285. 少数国家保持完整和最新的人口登记，可从中直接计算某些出生率和死亡率的分母。关于人口登记的论述见下文E节。

#### 1. 承受风险

286. 人口测度的基本原则是，任何一种事件的数量都必须与“承受风险”相联系。通常公认的原则是将事件数量除以相应的“存活人年”数或“承受风险”——一个把可能经历过事件（“承受过”事件“风险”）的人数与各人承受风险的时间长度合在一起的统计量。得出的系数是一个比率，更确切地说是事件的“发生-承受风险率”。

287. 承受风险原则对计算出生率和死亡率有两个基本影响。第一，计入任何出生率分子的每一个出生事件都应当已经发生在列入该出生率分母的某个妇女的身上，对于死亡率也同样如此。第二，列入任何出生率分母的某个妇女的每一次生育都应当计入该出生率的分子，对于死亡率也同样如此。下面这一节给出的是这些条件没有得到满足的一个实例。

#### 2. 按常住地址分类

288. 为计算出生率和死亡率目的而对民事登记系统的出生和死亡数据进行制表时，一般应当按常住地址而不是按发生地点或登记地点进行地理分类。从中得出出生率和死亡率分母的人口普查和调查数据几乎总是使用根据常住地址点进行的地理分类。承受风险的原则规定，用于列入任何比率分母的人员的地理分类也同样用于计入该比率分子的出生或死亡数。

289. 为计算婴儿死亡率而对婴儿死亡数进行制表时，以及为计算孕产死亡率目的而对产妇死亡数进行制表时，这一规则存在例外。由于这些比率的分母是出生数，而不是人数，所以既可按发生地点制表，也可以按居住地点制表。婴儿死亡率和产妇死亡率既可按居住地点计算，也可按发生地点

点计算。

290. 一个简单的例子证明了如果按发生地点而不是按居住地点分类可能发生的这种误差。假设某个地方登记区内有一家医院，而其周边地区没有医院，住在周边地区的妇女长途跋涉地来到这家医院分娩。如果出生事件按发生地点分类的话，那么拥有医院的地区的许多出生事件是属于住在周边地区的妇女的事件。结果，对拥有医院的地区计算出来的出生率将会过高，而对周边地区计算出来的出生率就会过低。

291. 从第三章A节第3小节中所讨论的覆盖面误差来看，在使用按发生地点计算的出生率代替按母亲常住地址计算得出的出生率时，既可能导致漏计，也可能导致不当计入。住在A区但在B区的医院分娩的妇女的生育事件会被排除在为A区编制的表格之外，成为一个漏计的事件，但会被列入B区的表格，成为一个不当计入的事件。

292. “常住地址”的统计定义可能涉及一些复杂的问题（联合国，1998年a，第2.20–2.24段）。《关于生命统计系统的原则和建议，第2次修订本》（联合国，2001年b）建议民事登记系统采纳人口普查中使用的常住地址的定义（同上，第276段）。

### 3. 估计分母

293. 几乎始终有必要根据人口普查和调查数据估计出生率和死亡率的分母，因为这些数据不直接提供这些分母。人口普查数据一般每隔10年可获得一次，各普查间隔期间进行的大规模样本调查通常不会超过一次。但是，分母一般需要用于每一个历年的分母，也常常需要用于更短的期间。

294. 从定义上讲，出生率和死亡率的分母是指承受风险的“人年”数。实际上，一个人口组在一个期间中的存活人年数常常近似于该组在该期间的中点时的人数乘以该期间的长度。下面的讨论将假设使用的是这一近似值。

295. 有三种估计方法：内插法和外推法、人口方程式法以及模型法。内插法和外推法比较简单，所需数据较少，适用性最广泛，但是，当人口变化没有遵循数学程序中假设的模式时，它们给出的结果可能不能满足要求。人口方程式法能够给出

非常精确的结果，但是这种方法比较复杂，所需的数据常常不能获得。模型法有时能给出最佳结果，其中有些方法广泛适用，但是计算过程比前面两种方法往往复杂得多。

296. 为了说明这前两种方法，考虑一下估计(a) 总人口和(b) 某一年龄组在没有进行普查的某个时点的总人数的问题。

#### (a) 内插法和外推法

297. 如果总人口规模的普查数字可从特定年份前和特定年份后的普查中获得，那么估计人数可通过两个年份间的内插获得。常常采用使用普查间距人口增长率的指数插值，但也可以使用其他内插法。

298. 如果只能获得某个前一年度的人口普查数字，则利用一个估计的人口增长率从普查数字推断估计人口。如果人口增长率大致保持稳定，则可使用最近获得的普查间距人口增长率。如果人口增长率一直在变动，则可为变动留有一定的容许差。

299. 这一程序也可用于内插或外推某个特定年龄组的人数。在这种情况下，得出的估计值的质量可能严重受到普查或调查年龄分布的误差的影响。在年龄分布误差严重的情况下，应当作适当调整。

#### (b) 人口方程式法

300. 人口方程式也称为平衡方程式，即

末端人口 = 开端人口 + 增加的人口 - 减少的人口

它适用于在任何期间被观察的人口。如果所述人口中没有经历任何国际移徙事件，则增加的人口为出生数，减少的人口为死亡数。如果存在国际移徙事件，则增加的人口为出生数与迁入移民之和，减少的人口为死亡数与迁出移民之和。

301. 人口方程式法利用民事登记中的出生和死亡数据，需要时加上国际移徙的数据和/或估计值，来估计某次普查或调查时间与晚些时候或早些时候的间隔期间的人口变动（增加数减去减少数），最好对晚些时候或早些时候的人口作出估计，因为短于历月的期间的民事登记数据通常不可能获得，所以通常需要对出生数和死亡数作些内插。



302. 在对两次人口普查间隔期间的某个时间估计人口数字的时候，可根据前一次普查作出“前向”估计，根据后一次普查作出“后向”估计。由于所获得的数据永远不是确切的数据，因此在前向估计与后向估计之间将会存在一些差异。这种差异对估计误差大小是一种有用的指示值。

303. 如果净国际移徙可以忽略不计或能够得到适当估计时，人口方程式法对国家一级的总人口估计常常起到有效的作用。它对总人口的国以下各级估计不常起到有效的作用。国内移徙对国以下各级地区、尤其对小的地区来说可能是人口变动的主要组成部分，国内移民人数数据常常难以得到。

304. 人口方程式法在许多情况下是不切实际的，因为没有得到有关人口的增加和减少的必要数据。例如在估计特定年龄组的人数时，增加的人口包括在所述期间达到该年龄的年龄下限的人，减少的人口包括在所述期间达到该年龄组的年龄上限的人。这些数字很少直接获得，它们的估计可能成为问题。

#### (c) 基于年龄组的估计与基于同期出生群的估计

305. 在估计特定年龄组的人数时，据以作出估计的普查数据或者可能是各年龄组的人数，或者可能是与特定年龄组对应的同期出生群的人数。例如，假设将15-19岁的妇女人数内插至两次十年进行一次的人口普查之间的时间中点。按年龄组计算就插入第一次普查时15-19岁的妇女人数与第二次普查时15-19岁的妇女人数之间。按同期出生群计算就插入第一次普查时10-14岁的妇女人数与第二次普查时20-24岁的妇女人数之间。

306. 在年龄报告比较准确的情况下，假如人口年龄分布存在严重不规则现象的话，则按同期出生群计算将比按年龄组计算更加准确。在估计按一岁分组的人数时，情况尤其如此。如果年龄报告质量不好，一般就应当使用按年龄组计算的办法。

#### (d) 模型法

307. 上述两种方法原则上比较简单，尽管应用时常常需要认真注意许多小的细节。估计年龄别出生率和死亡率分母的现有方法要复杂得多。这些方法可能涉及内插/外推与人口方程式法的某种结

合。实例见Coale（1984年）和Stupp（1988年）论著中所述的方法。

308. 人口预测法所具有的优势是，提供对所有年龄组 and 所有期间都相互一致并与估计的年龄别生育率和死亡率相一致的估计分母。在人口预测法中，对改变年龄分布的历史进行了重建，以便完全适用于可获得的数据。使用任何其他方法计算出来的分母在这一方面一般不会一致，尽管不一致的情况在统计上可以忽略不计。

#### 4. 国家一级的出生率和死亡率

309. 根据民事登记数据计算国家一级的出生率和死亡率时，只需将获自适当表格的出生数或死亡数（见上文A节第6小节）除以根据人口普查和户口调查估计的相应分母（见前一节）。

310. 在一些国家，公民中有相当多的人在外国生育或死亡。民事登记系统可以为登记发生在国外的相关事件作好准备。如果居住在国外的公民的事件被列入比率的分子，则承受风险的原则要求将这些公民也计入比率的分子。在估计分母时应当考虑到这一点。

311. 在一些国家，非本国公民的生育事件数也可能相当大。非本国公民可能包括旅居者、居住在该国的外籍工人或难民。如果这几类人中任何一类人向民事登记系统进行了登记的话，就需要对计算比率时对这些人的作何处理的问题作出决定。根据承受风险原则，如果将他们列入出生和死亡表和比率的分子的话，就应当将相应的人口数列入分母。

312. 国家惯例各不相同，但是许多国家的民事登记将登记发生在旅居者和外籍工人身上的事件，但不登记发生在难民身上的事件。然而，旅居者事件有可能被排除在出生和死亡表之外，因为人口普查通常不会查点旅居者。外籍工人的出生数和死亡数可能被列入表格，但是与公民的出生数分开。

313. 按一岁分组的年龄别出生率和死亡率需要用于某些目的。生育率和死亡率风险的变动率对某些年龄来说通常很高。当比率只是按五岁年龄组提供时，这一变动的细节就会被掩盖。与此同时，计算按一岁分组的准确比率时，既需要民事登记系

统的准确年龄报告，也需要人口普查或据以计算分母的其他来源的准确年龄报告。

#### 5. 国以下各级地理区域的出生率和死亡率

314. 如上文A节6(c)小节所指出，为基本所有国以下各级地理区域提供数据的可能性是作为一种生育率和死亡率数据收集方法的民事登记的最重要优势之一。从原则上看，计算国以下各级地理区域的比率的过程与计算整个国家的比率的过程并无不同之处，但是在实践中，常常存在重大差异。

315. 与估计整个国家的比率分母相比，估计国以下各级地理区域的分母几乎总是更加困难一些，尤其是对较小地区来说。内插法和外推法始终可以使用，但是较小地区的年度净移民数字的波动可能会使这些数字的准确性不如在国家一级使用的同样方法。使用内插法和外推法的期间越长，这个问题就越严重。上文B节第3(d)小节中所讨论的人口方程式和其他更加先进的方法常常是不适用的，因为关于国内移民的必要数据不可能获得。

316. 计算国以下各级地区的比率时要求：民事登记系统和人口普查在提到常住地址时必须使用同一个定义。它还要求在民事登记和人口普查中对旅居者、外籍工人和难民的处理进行协调。例如，如果旅居者的出生事件和死亡事件得以登记而普查时没查点旅居者的话，则民事登记系统应当提供不包括旅居者出生数和死亡数的数据表。

317. 计算极小地区的比率时引起的另一个问题是，计入比率分子的事件数量可能太小，无法提供可靠的信息。评估由于人口数量小而引起的随机变异性的方法与评估抽样误差的方法相似。对这一问题的详细分析不属于本手册的范围，但是一项有用的经验规则是：如果分子（出生数或死亡数）小于25，就不应当计算比率。

318. 过去，小于主要民政区的地区的出生率和死亡率通常是不列示的。然而，有了当代的信息技术，完全可以为大量极小的地区制表和储存大型出生和死亡表并计算比率，这种地区有几万、几十万甚至几百万，而且人口规模很大。过去几十年地理信息系统的发展为这类极其详尽的空间数据的分

析提供了概念和工具。关于强调普查适用范围的介绍，见《地理信息系统和数字化地图手册》（联合国，2000年b）。

#### 6. 人口亚群的出生率和死亡率

319. 人口亚群用诸如性别、种族血统、城乡住所和出生地等特征界定。某些特征，如城乡住所，可能会随时间而变动。其他特征，如出生地，按照定义在个人生存期中是固定的。还有一些特征，例如种族血统，在时间上是含糊不清的。从某种意义上讲，种族血统是不随时间而变化的。不过，在人口普查中，种族血统几乎始终是自动报告的，个人可能会随着时间而改变其自动报告的方式。

320. 即使某个特征按照定义是一个人终身不变的，但它在一个人的一生中可能在不同的时间有不同的报告。就拿出生地来说，它在一个人年轻时可能报告是正确的，但当这个人年老时，就会把它误报为当前居住地。

321. 计算人口亚群的比率的方法在许多方面与计算国以下各级地区的比率相似。必须将来自民事登记系统的提供比率分子的出生和死亡表与来自人口普查和户口调查的提供（通过估计）比率分母的个人情况表配对。这需要对两种数据来源的统计定义进行协调。

322. 在计算人口亚群比率过程中引起了一个问题，但在计算国以下各级地区的比率时不会出现这个问题。关于民事登记系统中的出生和死亡的特征或人口普查和户口调查中的个人的特征的信息有可能漏报。例如，按种族血统分列的出生事件表可能会包括一个“未报告种族血统”类别。处理这种未报告的数量通常方法是对它们按比例分配。但是，这将会引起某种误差，如果未报告的百分比比较高的话，误差可能会相当之大。如果民事登记和人口普查的数字中都有未报告数字的话，问题就会变得复杂化，情况通常就是这样，因为未报告的模式在两种来源之间可能是不同的。

323. 人口亚群的比率分母的估计不同于对国以下各级地区的估计，因为它可能涉及对亚群之间的人口流动、而不是常住地址的变动的估计。就

随时间变动的特征而言，通常必须使用内插法和外推法，这是因为其他估计方法所需的数据一般不存在。

### C. 登记完整性的概念

324. 登记完整性的标准定义指事件（出生和死亡）和那些发生在特定期限——通常指一个历年——的事件的登记情况。某个特定人口（国家人口、国以下各级地区或人口亚群）在某个特定期间的出生登记的“期间”完整性界定为：在该期间内进行登记的出生数占总人口在该期间内所发生的出生事件数的百分比。死亡登记的期间完整性的定义与此相似。

325. 在计算期间完整性时，可以为事件的发生与登记之间的容许最大延迟时间规定一个容许差。例如，假设登记法规定死亡事件必须在事件发生14天之内予以登记。那么，历年2000年的死亡登记完整率的分子就可能包含2001年头两个星期内登记的死亡事件以及2000年期间登记的死亡事件。

326. 登记完整性还可界定为某一特定年度内发生的并已在该年度内登记的或在发生后的规定时间内登记的事件数。理想的方法是强行做些删节以免涉及不确定的将来。例如，完整性可以表示为自事件发生年年底算起10年内登记的事件的百分比。

327. 这种测度可以称为“同期出生群”完整性统计。就出生事件而言，它们指的是在某个特定历年或其他期间出生的同期群的后来的登记史。就死亡事件而言，它们指的是在某个期间内死亡的同时出生群，这是一个不太常见的概念，但是适用于对死亡事件延迟登记的研究。关于延迟登记的讨论，见下文C节第4小节。

#### 1. 未登记出生和死亡事件的组成成分

328. 有时不妨把民事登记系统没有运作的地区所发生的、因此没有可能进行登记的未登记出生和死亡事件，与民事登记系统已经运作但是没有登记全部事件的地区所发生的出生和死亡事件区别开来，这样做是有用的。

329. 这种概念区分的一个简单而重要的用途是承认民事登记系统提供的出生数和死亡数在地

理信息方面应当足够的详尽，以便能够区分这两类地区。然后可以对民事登记已经运作但是登记不完整的地区的登记完整性作出估计。这些估计对估计整个国家的登记完整性可以起到一种有用的辅助作用。

330. 民事登记系统在城市地区和一些农村地区运作而在另一些农村地区没有运作的情况并不常见。

#### 2. 选择性登记

331. 当登记的完整性在各地理区域之间、人口亚群之间或年龄、性别等其他人口特征之间存在差异时，就发生选择性登记。即使不存在选择性不完整登记的话，也仍然必须对出生和死亡事件的登记数作出调整以便删节，但是只需一个调整系数即可以用于所有出生和死亡事件。

332. 应当对选择性登记尽可能作出分析和评估，以确保民事登记数据不至于传递一种扭曲的人口条件情况。例如，经常会出现这样一种情况：出生和死亡事件的登记完整性在城市地区高于农村地区。如果把一个对整个国家来说是准确的调整系数分别用于城市地区和农村地区，那么在这种情况下就会夸大城市地区的出生和死亡水平，而缩小农村地区的出生和死亡水平。

333. 大多数估计登记完整性的方法都作如下假设：在某个人口亚群内，某些特征如年龄方面都一致存在登记不足的情况。在评价这些方法产生的完整性估计时，关于这些特征方面的选择性登记的信息是有用的。

334. 选择性登记有时可带来预料之外的、意想不到的结果。例如，一份研究报告发现，某些有大学公共卫生学校的美国城市的死亡率高于没有这类学校的城市。该研究报告总结说，公共卫生学校实际上是造成这种差异的原因。然而，成为原因的因素是这些学校给登记完整性所带来的积极影响——而不是对卫生条件的消极影响。

#### 3. 发生日期和登记日期

335. 出生或死亡事件的登记日期应当始终是民事登记系统有关事件所收集的信息的一部分。登

记日期是登记过程的一个人为现象，而不是事件本身的一个特征，但重要的是应当评估登记延迟的程度。

336. 当出生或死亡事件直到民事登记法规定的最后期限以后才予以登记时，就发生延迟登记。例如，有可能大部分出生事件直到儿童到达上小学的年龄时才得到登记，而登记的目的是为了儿童得以入学。

337. 与出生事件的延迟登记相比，死亡事件的延迟登记可能不太常见，原因很简单，因为死者不再存在于社会。然而，评价民事登记时应当详细审查延迟死亡登记以及延迟出生登记的证据。

#### 4. 延迟登记的分析

338. 延迟登记是导致某个特定年度期间或其他期间不完整登记的一个可能原因。对出生或死亡事件按登记日期而不是按发生日期进行制表时，它还可能导致死亡数字过高，详情见上文A节第6(a)小节。

339. 为了评估延迟登记的程度，民事登记系统可以每年产生显示按发生年份分列的本年登记的所有出生和死亡事件的表格。表格可以包括其他一些特征，例如，关于出生事件可包括出生时母亲的年龄，关于死亡事件可包括死者的性别和年龄。

340. 这些表格可提供有关延迟登记程度的宝贵信息。如果延迟登记程度不严重，那么任何特定年份登记的几乎所有出生或死亡事件都应该是在同一年份发生的。如果有许多出生和死亡事件的登记被延迟了，那么任何特定年份登记的事件中有许多应该是在前几年发生的。表格的补充范围将根据相应的特征揭示各次登记延迟之间的相同之处和不同之处。登记延迟之间的不同之处是导致选择性不完整登记的一个可能原因。

341. 累积了两年或两年以上的表格特别宝贵。例如，假设按发生年份编制的各年登记的出生表可提供1990–2000年的表格。1990年的表格将显示1990年登记并发生于1990年的出生数。1991年的表格将显示1991年登记但发生于1990年的出生数，后面的年份也同样如此。年度系列表便于跟踪任何年

份出生的同期出生群，例如1990年出生的同期出生群，并便于了解多少人是在那一年出生的，多少人是在下一年登记的。

342. 当延迟登记的模式比较正规时，它们可用于根据后来几年发生和登记的事件数估计任何特定年份发生的事件数。这提供了一种估计由于延迟登记而导致的登记不完整性的方法。生命表法可用于分析某次事件的发生与民事登记系统对事件的登记之间的延误，使用的方法大体上与用于分析寿命长度时所使用的方法相同，寿命长度可被认为是出生与死亡之间的一种“时间间隔”。关于延迟结婚登记的实例，见Feeney 和Saito的论著（1985年）。

#### 5. 登记完整性的趋势

343. 评估登记完整性时，一般应当假设完整性是随时间变化的。因此，登记完整性的说明必须提及特定年份或其他期间。

344. 对每一个历年直接估计出生和死亡登记的完整性往往是不可能。每十年中可能只进行一次或两次评估完整性的专门调查。通过利用人口普查或调查数据估计完整性的方法，可以得出比历年长得多的期间的估计，例如两次普查间隔期间。

345. 因此，评估登记完整性变化时有时采取质量评估的形式：不论完整性提高还是下降，不论是逐步性变化还是突然性变化，如果是逐步性变化，那么每年大致变动率是多少。

346. 如果登记完整性提高速度较慢，可望出现逐步性变化。导致突然性变化可能有各种原因，例如由于民事登记系统的范围被扩大至以前没有提供服务的区域，由于开展了目的是劝说公民登记出生和死亡事件，或者由于宣布改变延迟登记的罚款额。民事登记数据年度报告中提供的文献资料可以提供有关可能影响了登记完整性的活动、变化或发展的信息。

347. 登记完整性的逐步性变化通常可用十年或十年以上期间的线性趋势作粗略估计。完整性的线性变化假设有时用于为登记完整性变化建模并对它作出估计（Macura, 1972年）。

#### D. 不完整民事登记数据的用途

348. 不完整民事登记数据的一般用途是估计登记完整性的水平，并用这一估计对事件的登记数字作向上调整以估计真实数字。这项工作一般以一个期间为基础，方法是将某一特定年份或其他期间发生的已登记出生和死亡事件的数字乘以一个校正系数。校正系数只是估计的登记完整度的反数。

349. 当把调整系数用于按年龄或其他特征分列的出生或死亡的分布时，调整后的分布的准确性同时取决于登记完整性估计的准确率和选择性登记（上文C节第3小节）。当把调整系数用于连续几年或其他期间的数据时，必须把登记完整性的变化考虑进去（上文C节第5小节）。

350. 出生和死亡登记完整性的系统评估应当作为每个民事登记系统的正常运作的一个组成部分。如果登记严重不完整，则完整性估计对于使数据有用于统计目的就毫无价值。如果登记完整或接近完整，则必须作出估计以确保登记的完整性不会随影响登记系统的条件的改变而恶化。

351. 第三章A节第3小节提到了作为一项评价数据质量的一般方法的记录一配对的研究报告。为了应用于民事登记数据，这些研究把各个出生或死亡事件的民事登记记录与获自人口普查、户口调查或一些其他来源的记录进行配对。这类研究报告如果可以提供，应当加以利用。但是，这类报告常常不可能获得，在这种情况下，登记完整性估计可以通过与人口普查和户口调查数据的总体比较进行，详情见下文D节第2小节。

##### 1. 不完整登记的指示值和估计

352. 出生和死亡事件的不完整登记的指示值就是表明某些事件没有登记、但是并不提供不完整登记程度的数量估计的统计量。如果没有获得完整性的估计，则指示值对于产生估计值的指导性工作可能有用。

353. 不完整登记的最简单指示值是过低的出生率和死亡率。例如，如果从全国人口中观察到的出生预期寿命最高水平为大约80岁，那么获自得出预期寿命为90岁的民事登记的一套死亡率数据几乎肯定反映了不完整死亡登记。

354. 其他指示值以简单的比较为根据。例如，如果农村死亡率大大低于城市死亡率，就有可能农村的死亡登记不完整。这个推论的根据部分可能是获得农村地区的完整登记更加困难，部分可能是城市地区的死亡率可能低于农村地区。

355. 民事登记提供的空间和时间详情可作为更加复杂的比较的根据。下文援引的是美利坚合众国的一份早期登记完整性研究报告：

如果某地区的死亡率上升而所有毗连地区的死亡率下降，就有根据怀疑上升是由于记录得到改进的缘故。……如果中央地区上升幅度较大而周边地区上升幅度不大，也是同样的道理。如果城市附近的农村死亡率极低而毗连城市的死亡高得多，如果国家的低死亡率一直上升而该城市没有上升，就表明该国各地区的死亡记录一开始不完整，但后来一直在不断改进（Willcox, 1940年，第205段）。

诸如上述这类观察资料并没有为调整登记的出生和死亡数字提供任何根据，但是它们确实提供了有关登记完整性的有用信息。

##### 2. 与普查和调查数据的总体比较

356. 总体比较的一般方法是确定可从现有的民事登记数据和另一个可能更加准确的来源提供的的数据中计算出一个统计量。如果按正确方向去做，这两种统计量之间的差额就可以归于不完整登记。登记的完整性的估计方法是：确定必须对事件的登记数使用多大的调整系数以使从民事登记中推算出来的统计量与从其他来源推算出来的统计量相一致。

例1. 假设一项生育率调查提供的调查访问前三年的总和生育率为每个妇女3.4个子女，根据民事登记数据计算出来的同期总生育率为每个妇女2个子女。把民事登记数值乘以 $3.4/2=1.700$ 将得出调查率。默示的登记完整率是 $2/3.4=0.600$ ，或60%。注意：以比例表示的登记完整率是必须适用于登记的事件数以获得估计的真实事件数的调整系数的倒数。

357. 在实践中，计算方法没有如此简单，但原理是相同的。生育率调查访问一般持续进行好几

个月，所获得的信息通常涉及访问日期，而不是任何固定基准时间。提供的民事登记数据一般是日历期间、日历年、日历季度或月数据，这些时间通常与调查中的实地调查期间不完全一致。进行比较时需要对这些信息来源的时间基准作适当调节。

例2. 假设出生预期寿命是根据普查数据或调查数据估计出来的，为68岁，根据民事登记数据计算出来的预期寿命为75岁。显然，把民事登记死亡率乘以某个大于1的系数并重新计算生命表将会得出一个较低预期寿命，但是没有简单的公式可以计算出将会使民事登记的预期寿命从75岁减低到68岁的系数。然而，借助现代化个人电脑的计算能力，便可以轻而易举地使用计算方法获得理想的系数。

例2与例1的根本区别是所比较的统计量与从中计算出统计量的年龄别生育率之间缺乏一种线性关系。把一组年龄别生育率乘以任何系数将会使从中计算出来的乘以同一系数的生育率成倍地增加。然而，把一组年龄别死亡率乘以某个系数将会使乘以另一个系数所得出的预期寿命成倍地增加。

358. 计算机空白表格程序提供了一种解决找到理想系数问题的简单方法。第一个步骤是研制出一种空白表格，该表格把年龄别死亡率作为输入信息并产生一份包括出生预期寿命值在内的生命表作为输出信息。这种空白表格可加以编辑，以便使输入的年龄别死亡率能够被空白表格的一个单元格中的一个系数相乘。于是这个单元格中的数值被定为给出理想的出生预期寿命的数值。这一计算过程可以通过尝试法或利用为此目的设计的空白表格功能完成。

359. 在前两例子中，比较的统计量是一个常用的人口统计指标，从其本身来说具有重要的价值。不过，正如下面的例子所述，也许选择作为一个不熟悉的人口统计指标的统计量的方法会使这种比较变得更加容易。

例3. 假设可以获得间隔期为五年的两次普查间隔期间每一年的按一岁分组的出生时母亲年龄分布的登记出生数。还假设一个关于妇女平均生育子女数的问题已被列入这两次普查，并可以获得两次普查的按一岁分组的妇女平均生育子女数。第一次普查所获得的妇女在任何特

定年龄的平均生育子女数与第二次普查获得的妇女在这个年龄加上五年时的平均生育子女数之差即为两个年龄之间的累计同期群生育率估计。可以从出生登记数据中推导出相同统计量的估计，然后可将它与普查导出的估计值相比较，以估计对特定同期群和年龄组的报告完整性。有关这一程序的例子见Coale、Cho和Goldman的论著（1980年，第14-18页）。

在例3中，所比较的统计量作为一个生育率的测度并不常见而且意义有限。然而，一旦比较得以完成，所得到的登记完整性估计便可用于普查间隔期间登记的出生率，结果可用于根据不完整民事登记数据估计标准的年龄别出生率。

### 3. 利用模式

360. 民事登记数据的某些使用并不取决于登记是否完整，只是取决于随着时间的流逝和在某些人口群体中大致保持不变的不完整率的程度。

361. 例如，如果登记完整性大致保持稳定不变，根据民事登记计算出来的出生率和死亡率趋势即使在水平较低的情况下也是正确的。除了根据易受各种不同误差影响的其他来源得出的估计数外，这可能也是重要的信息。例如，用亲生子女法或根据生育史估计的出生率可能提供一个更好的总体水平指示值，但是由于年龄或日期的误报，也可能提供一种扭曲的趋势指示值。在这种情况下，生育率的总体水平可从一种来源确定，趋势可从另一种来源确定。关于使用这种方法进行分析的例子，见Retherford、Mishra和Prakasam的论著（2000年）。

362. 同样，根据民事登记得出的出生或死亡的年龄分布即使在登记水平较低的情况下也可能大致是正确的。例如，根据不完整民事登记数据计算出来的年龄分布可用于根据人口普查或调查进行的生育率估计。

363. 死亡事件年龄分布可以给出一个有用的成年人死亡率的指示值，虽然婴儿和儿童死亡报告的完整性水平可能不同于成年人死亡报告的完整性水平。尤其是当艾滋病病毒/艾滋病发病率较高时，成年人死亡事件的年龄公布显然表明存在艾滋病死亡事件，在存在艾滋病死亡事件的年龄范围内，死

于其他原因的死亡率相对较低。

#### E. 人口登记作为生育率和死亡率数据来源

364. 人口登记是一个不断更新的数据集，其中含有人口中每一个人的一份记录。充分完善的人口登记系统可能成为生育率和死亡率数据的一个理想的信息来源。人口登记是一个组织各种数据收集系统所收集的数据的方法，其中包括收集出生率和死亡率数据的民事登记系统，和经过密切协调的收集有关迁入移民、迁出移民、居住地点的变动和诸如受教育程度或职业等个人特征变化的信息的系统。许多国家保持某种人口登记，但是只有极少数国家的人口登记已达到充分完善的水平，以致完全可以提供可与本手册中讨论的其他来源提供的数据相比的生育率和死亡率数据。

##### 1. 概念

365. 如本章B节中所述，根据民事登记数据计算出出生率和死亡率是把出生和死亡数除以相应的人数。由于利用一种数据来源（民事登记）提供分子，利用另一种来源（人口普查和调查）提供分母，引起计算过程相对复杂。

366. 民事登记在某些情况下除了可用于提供出生率和死亡率的分母外，还用于提供分子。假设一个国家为人口中任何一个人的一生建立了完整的出生和死亡登记，时间长达约100年。还假设该国在这个期间没有任何迁入移民和迁出移民。最后假设民事登记系统每次登记一个出生事件时，便为出生的这个人建立了记录，民事登记系统每次登记一个死亡事件时，便找到了死者的记录并在上面记录了死亡的事实。

367. 从原则上讲，用这种方法建立的个人记录可以提供关于每一个时点的人口规模和年龄-性别分布的完整信息。要确定例如在某日午夜时分的人口数时，只需确定对在此时之前出生和在此时之后死亡的人的记录数。为了确定某个特定年龄组的这些人的数量，只须选择那些有相应出生日期的记录。

##### 2. 登记迁入移民和迁出移民

368. 各国人口，特别是在长期内，始终都

在经历迁入移民和迁出移民事件。按照国际承认的定义，迁入和迁出事件并非属于生命事件，尽管它们对人口变动像出生和死亡事件一样“重要”。因此，民事登记并不收集关于国际移民的人数和特征的数据。

369. 因此，除了民事登记系统外，保持人口登记还需要一个登记进入一个国家的每一个迁入移民和每个离开这个国家的迁出移民的系统。如果这样一个系统已经存在或可以建立，而且提供有关每一个国际移民的完整和准确的信息，它就可以用来更新国际移民的人口登记，使用的方法可与用来更新自然人口变动登记册的民事登记记录的方法大致相同。

370. 每次对迁入移民进行登记时，都将对人口登记册进行搜索以确定是否留有这个人的记录，因为有可能这个人以前是该国人口的一个成员。如果找到了有关这个人的记录，将利用有关最近一次迁入时间的信息进行修订。

371. 如果人口登记册上没有关于该移民的记录，将建立一份新的记录，将该迁入移民的性别和出生日期以及迁入日期登记上去。

372. 每次对迁出移民进行登记时，将在人口登记中找到迁出者的记录并进行更新，以反映迁出的详情。

##### 3. 登记居住地址的变化

373. 无论在国以下一级还是在国家一级，生育率和死亡率数据几乎始终是需要的。对于提供国以下一级信息的人口登记来说，它必须提供有关居住地址的信息。为了更新这一信息，需要有一个登记居住地址变化的系统。建立这样一个系统的难度可能大于建立一个民事登记系统或建立一个登记迁入移民和迁出移民的系统。这也许是人口登记没有被普遍用作生育率和死亡率数据来源的主要原因。

374. 如果登记居住地点变化的系统已经建立或可以建立，利用这一系统保持最新人口登记的程序与已经讨论过的程序相似。每当对居住地址变化进行登记时，在人口登记册上找到有关人的记录并进行更新。如果居住地址的变动得以登记，并且如果提供的信息是准确的话，则人口登记册可用来产生任何时

间、任何地区、任何年龄-性别组的人数。

#### 4. 登记其他个人特征的变化

375. 居住地址是随时间而变化的诸多个人特征之一，可以用来计算特定人口亚群的出生率和死亡率。人口登记提供人口亚群的出生率和死亡率数据的能力受到个人记录上所包括的信息的限制。可能随时间变化的各种个人特征需要一个为人口中的每一个人更新登记册中的信息的系统。

376. 所需的信息可从行政记录中获得，例如教育机构保持的学生记录或就业保障系统保持的失业记录。人口登记原则上可以利用这一信息更新人口中的每一个人的受教育程度和就业状况信息。实际执行是另一个问题。这种方案目前只有在少数几个发达国家才是切实可行的。

#### 5. 建立并保持人口登记册

377. 建立人口登记需要建立含有有关人口中每一个人的一份记录的数据集。每次人口普查都根

据不完整的人口查点建立这样一个数据集。

378. 保持人口登记必须对记录不断进行更新，具体任务包括如下：

- 登记出生或迁入移民事件时必须建立新的个人记录
- 当人口中的某个人改变其居住地址时，必须对记录作相应更新，对人口登记册中的任何其他个人特征也同样必须作相应更新
- 当死亡或迁出移民事件发生时，必须对死亡者或迁出者的记录进行更新以反映有关事件的相关信息

379. 因此，为了用作生育率和死亡率数据的来源，人口登记必须得到数据收集系统的支持，数据收集系统则必须登记每一个移居该国的迁入移民、每一个离开该国的迁出移民、每一次居住地址的变动和可能随时间而变化并登入登记册中的其他个人特征的每一个变化。除了需要建立登记所有出生和死亡事件的民事登记系统以外，还需要满足上述要求。



## 五、来自普查和调查的生育率数据

### 导 言

380. 本章研究人口普查与抽样调查中可能包含的各种问题以产生生育率数据，重点论述有关年龄、与户主的关系和出生婴儿数和存活子女数的问题之间的相互作用。

381. D节谈到的出生婴儿数问题提供通常无法从包括完善的民事登记制度在内的任何其他来源获得的重要的生育率数据。出生婴儿数在世界范围内的普查和调查中使用非常广泛。把这个问题与第七章A节所述的存活子女数问题放在一起调查时，它能提供有关婴儿和儿童死亡率的重要可利用数据。

382. 通过普查和调查的方法用以产生生育率数据的最详细的问题是那些产生每个妇女的生育史的问题，主要是一份妇女的全部子女名单，列示每个子女的出生日期以及各种其他资料。F节所述的生育史是生育率调查的特征，但有时被用于大规模的户口调查，在极少情况下甚至被用于人口普查。说是“生育”史，却是关于婴儿和儿童死亡率的一个重要资料来源，因为它们几乎总是包括每个出生的子女在调查时是否存活的问题，如果没有存活，还包括死亡日期。

#### A. 存活倒推

383. 人口普查中点查的一岁以下婴儿的数目可以被看作是普查前年度中人口出生数目的估计数。如果可以忽视移徙的情况并且可以获得死亡率的估计数，那么可以通过向上调整人数以计入在进行普查前死亡的人数来提高估计数。这一调整被称为存活倒推。

384. 如果没有有关出生的完整的民事登记数据，一般可以用存活倒推来估计每次人口普查和某些或全部人口调查前15年的出生人数、粗出生率和一般生育率。更长期间的估计也是可能的，但出于

各种原因，也许会成问题，而且由于普查一般每十年进行一次，较长数列的价值通常很小。如果可以得到生育率年龄类型，同样可以估计总生育率。

385. 拥有来自连续的普查和/或调查的重叠的估计数数列非常有价值，因为这提供精确性测试。下一节说明协调从不同数据来源产生的生育率数据的重要性。与结合1990年进行的普查以及两次普查期间的10年中进行的一次或多次调查中得到的类似估计数得出的估计数相比较。仅仅从2000年进行的普查中得到的存活倒推估计数的价值则要小得多。

#### 1. 问题

386. 存活倒推所必需的惟一信息是按一岁分组的每一个普查或调查对象的年龄。该信息可以通过有关出生日期（优选）的问题或者直接询问年龄得到。不过，存活倒推需要有关死亡率水平的估计，这一信息常常可以通过有关出生子女和存活子女的问题提供。出生子女的问题见下文D节，存活子女的问题见第六章A节。

387. 对于衡量生育率和死亡率以及许多其他用途来说，年龄是所有人口特征中最为重要的。在西方文化中年龄观念既简单又根深蒂固。一个人在任何特定时间的确切年龄是自他出生后所流逝的时间。一个人的周岁年龄是低于他的实际年龄的最大整数（0、1、2……）。

388. 在人口普查和调查中收集年龄信息的优选方法是包括一个有关出生日期（年/月/日）的问题。比起直接询问年龄来，这有可能提供更准确的信息。在直接询问年龄时，年龄应始终以单岁的形式记录到至少97岁（98和99有时可以被归到“98岁或以上”和“不确定年龄”）。

389. 在收集各种人口普查和调查数据时要遵循的一个普遍原则是，尽可能使用被调查人最有可能给出精确回答的问题。就出生日期和年龄而言，

这意味着用日历制和被调查人和申报人使用的计算年龄的方法来问问题。例如，在许多亚洲社会，普查和调查问题应按照农历来询问出生日期。在数据处理过程中可以把农历日期转换为阳历的日期。不应利用调查工作者来做这样的转换，因为这一工作会把他们的注意力从更重要的收集尽可能准确的信息的任务上转移开，这样就得不到一致的准确结果。

390. 可以酌情加入更多的问题以便收集更准确的有关年龄和出生日期的信息。例如，在许多调查对象有诸如身份证或免疫接种记录卡等证件来表明他们的出生日期的情况下，调查可以包括一个关于是否可以向调查员提供这类证明的问题。

391. 在培训调查员的时候应始终强调出生日期和年龄的重要性，调查员清楚地向调查对象传达这一重要性常常是有必要的。在收集生育率数据方面，获取儿童和青少年的尽可能准确的出生日期或年龄尤为重要。

## 2. 制表

392. 每一次人口普查和大多数人口调查都会提供一张按性别和按一岁分组的总人口表。该表是在出生人数和出生率的存活倒推计算中使用的主要数据。

393. 即使在年龄报告质量差的情况下，也应得出并分析单岁年龄分布。按一岁分组的年龄分布可以更仔细地审查数据以查找误差，并且可以给年龄进行重新分组以尽可能消除年龄误报的影响。

## 3. 估计

394. 人口普查或调查前一年、两年……期间出生人数估计是通过把存活倒推因素运用于年龄为0、1、2、……岁的人数计算出来的。这些因素通常根据适用于计算估计数的年份期间的人口的估计生命表计算。如果死亡率发生变化，并且可以得到普查或调查前数年的一系列生命表，那么可以根据一系列生命表而不是单表计算。

395. 对许多国家来说，本章D节和第六章A节所探讨的出生子女和存活子女的问题对于这些计算中使用的估计生命表是一个重要来源。当出生

子女和存活子女数据被用于存活倒推计算时，直接计算存活倒推因素而不计算生命表的便利方法可见Brass的论著（1979年）。

396. 估计粗出生率和一般生育率分别需要普查或调查前各年年中时总人口数和育龄妇女人数。这些数字可以再次使用一个或多个估计生命表通过普查或调查年龄分布的存活倒推估计。

397. 上面阐述的存活倒推计算不考虑移徙情况。在国家一级，移徙情况可以忽略，但对于国以下区域，特别是小区域而言，移徙常常是一个重要问题。对时间更久远的出生人数和出生率来说，移徙的影响可能更大，因为一般说来，年龄较大的人移徙的可能性要大于年龄较小的人。尽管原则上有可能对移徙进行调整，但似乎在实践中很少这么做。例如，关于出生地点和/或居住时间长短的信息可以用于这一目的。

## 4. 质量

398. 对出生人数和出生率的存活倒推估计的质量取决于年龄分布数据的准确性、存活倒推因素的准确性以及移徙的范围（或者，如果对移徙作了调整的话，则包括调整的准确性）。

399. 年龄分布的误差对存活倒推估计的质量所产生的影响很容易理解。如果报告的年龄为0岁的人数少了10%，那么普查或调查前一年的出生人数估计也会少10%。误报年龄而产生的儿童和青年的年龄分布中的波动不定将表现为相应的一段时间内估计数的波动不定。然而，即使逐年估计非常糟糕，普查或调查前15年期间的出生人数和出生率估计也可以是比较准确的（Retherford、Mishra和Prakasam，2000年）。

400. 存活倒推因素中的误差所产生的影响也很容易理解，但在这种情况下，存在一种在年龄分布误差的情况下并不具有的顽固性。高估（低估）死亡率水平引起的存活倒推因素过高（过低）会导致出生人数和出生率的估计数过高（过低）。然而，在除高死亡率人口外的所有人口中，这些误差的影响被减弱了，因为存活倒推比率反映存活的比例而不是死亡的比例。

401. 例如，假设在某个同期群中有10%的出

生人口死亡，并且这一百分比的估计数低了30%，这是一个大误差。那么，估计的存活倒推比率为 $1/(1-0.07)=1.075$ ，相比之下，正确率为 $1/(1-0.1)=1.111$ 。估计的生存率当然过低，但其误差只有3.2% $(1-1.075/1.111)$ ，很少超过死亡率估计误差的1/10。

402. 如果不对移徙进行调整的话，移徙对存活倒推估计的影响很简单。如果由于移徙导致某个同期群失去（获得）成员，那么该同期群中出生人数的存活倒推估计将过低（过高）。如果对移徙作了调整，那么影响大概会减少，但具体影响将取决于具体的调整程序和所使用数据的准确性。

403. 在实践中，年龄分布数据的准确性通常是最重要的考虑事项。然而，在对非常小的国以下区域进行估计时，移徙水平高也许是一个重要的因素。存活倒推比率估计不准确所导致的误差一般是最不重要的因素，然而这取决于在死亡率水平和趋势估计中所采取的合理谨慎态度。年龄分布误差将在下一节详细讨论。

404. 在对国以下地理区域，例如城市和农村地区、州或省或级别较低的地理单位做存活倒推估计时，移徙会导致误差。尽管存活倒推提供普查或调查前15年或更长时间的追溯估计，但这些估计的基础是在普查或调查时的居住地。

405. 例如，研究根据2000年进行的一次普查对一国的城市地区所做的一个15年估计数列。许多在普查时居住在城市地区的人也许五年、十年或十五年前居住在农村，但他们的育龄期归于城市地区。对早些年城市生育率的估计将是对早些年居住在城市地区的人和居住在农村地区但后来移徙的人的生育率的合并估计。从农村迁到城市的移民的生育率低于农村非移民，这一趋势当然会削弱移徙的这一影响。

406. 评估存活倒推估计的质量的最有效方式是比较在不同的时点进行的各种数据收集活动所得出的估计数的时间序列。如果不同的趋势比较一致，那么这常常是估计数的质量还可以的一个好的（然而却是不确定）的迹象。如果不同的趋势差别迥异，那么这无疑表明一些或所有估计数存在误差。差异分析随后可以得出有关可用来提高估计准

确性的各个数列中的误差的大小和方向的结论。第三章A.4节的例子就是一个说明。这一方法的优点是解决估计中的所有误差。

## 5. 年龄分布误差

407. 年龄分布误差来自年龄错报和不同年龄组的人口漏查。年龄错报误差，尤其是那些可归因于年龄堆积的误差，是最常见的，但不一定是最重要的。为了评估年龄分布误差，考虑各种误差至关重要。

### (a) 年龄误报

408. 想象一个正方形的表，按真实年龄（行）和报告年龄（列）对有关人口进行分类显示，就可以对年龄错报有一个直观的了解。如果没有年龄误报，那么这张表只有从最上面的左边方框延伸到最下面的右边方框的对角线上才有非零项目。任何行中的非零方框的分布代表该行所表示的真实年龄的人的报告年龄的分布情况。

409. 由于年龄的误报，任何年龄或年龄组都可能增加和减少成员。如果只考虑所查点的人，那么所报告的任何年龄的人数 $x$ 将是真实年龄是 $x$ 的人的人数加上由于误报而从其他年龄转到这一年龄的人的净数。由于转来的净数可能为零，即使在出现大量年龄错报的情况下，年龄误报的程度也不同于被报告为某一年龄的人的人数与这个年龄的实际人数之间的差异程度。

410. 超过实际人数的人被报告为年龄以“0”和“5”结束的年龄堆积是误报的最常见和最不重要的形式。在其共同的较为明显的表现中，它立刻就能显现出来。存在显示其严重程度的各种指数并且一般要计算这些指数，但其用处不大。更重要的信息可以仅仅通过用曲线表示年龄分布得到。一般说来，年龄堆积对较大的年龄的影响超过对较小的年龄的影响，可以通过适当的调整相当有效地抵消其影响。

411. 在现述情况下年龄误报的更重要的形式是：(a) 把低龄儿童的年龄上提，(b) 以婚姻状况或生育子女的数目为依据影响年轻妇女的报告年龄，或者(b) 夸大老年人的年龄。

412. 对低龄儿童（5岁以下婴幼儿）的年龄的报告常常过高，导致低龄群体中的人的净减少。为什么会发生这种情况还不清楚，但这一误差对实践有重大影响。这导致对普查或调查之前一个阶段的估计过低，因而导致出现生育率下降，而实际并非如此，或者扩大现在的下降趋势的情况。在根据两次或多次不同的普查和/或调查作出估计时，很容易认识到这样的影响。第三章图1提供了一个突出的、有启发性的例子。

413. 年轻妇女的报告年龄也许根据她们结婚与否和有几个孩子而向不同的方向发生偏差。在年龄不是一个重要的社会考虑事项的情况下，调查对象也许不清楚自己的年龄，那么调查对象和/或调查员就不得不猜一个大概的年龄。在这种情况下，没有结婚或没有子女的妇女的报告年龄也许低于她们的真实年龄，已婚并有子女的妇女的报告年龄也许高于她们的真实年龄。这类误报可以导致10-14、15-19以及20-24年龄组之间的大规模的净转移。出于各种不明原因，男子也可能出现这样的转移。

414. 老年人有时包括年轻成年人的年龄可能被夸大，也就是说，报告年龄大于其真实年龄的人的比例一般会超过报告年龄小于其真实年龄的人的比例。这也许是因为某种观点认为高龄是一种个人的和社会的特征。Retherford和Mirza（1982年）的论文提供了一个有用的个案研究。

415. 三种年龄误报的误差比年龄堆积的误差更难评估，也许是因为这个原因，对它们的研究也较少，理解也不那么透彻。对它们的评估一般必须结合对年龄选择性人口漏查的评估，正如Luther、Dhanasakdi和Arnold(1986年)的研究报告中所阐明的那样。

#### (b) 年龄选择性人口漏查

416. 年龄数据的质量反映年龄选择性人口漏查以及年龄误报的情况。年龄误报扭曲被调查人口的报告年龄分布。年龄选择性人口漏查扭曲被调查人口的真实年龄分布。两种误差结合在一起就造成了(a) 被调查人口的报告年龄分布与(b) “真实”人口的真实年龄分布之间的差异，真实人口的意思是如果没有年龄误报和覆盖范围误差的话会被调查到的人口。

417. 在普查和调查中漏掉婴儿的趋势是年龄选择性人口漏查的最简单的例子，对存活倒推估计的准确性的影响也最大。例如，如果20%的0岁的人在普查或调查中被遗漏，那么普查或调查前的一年的出生人数估计将因此低20%。在实践中，很难区分年龄分布数据中年龄错报和年龄选择性人口漏查的影响。

418. 在只向某些年龄的人询问某些问题的情况下，年龄选择性人口漏查和年龄误报可以相互影响。例如，假设年龄在50岁以下的妇女被问及一系列非常详细的有关她们所生的全部子女的问题，而年龄在50或50以上的妇女不被问及这些问题，这就会促使那些宁愿尽量减少工作而不是获得准确信息的普查员多报接近50岁的妇女的年龄。

419. Ewbank（1981年）对普查和调查中的年龄分布误差作了较古老但非常详尽的探讨。

#### B. 亲生子女法

420. 亲生子女法是存活倒推法的延伸，即把用存活倒推法倒推出来的孩子尽可能地与他们的母亲进行配对。把子女与母亲配对可以把母亲的特征与她们的子女联系起来。这样一来就可以按出生时母亲的年龄分解用存活倒推法倒推出来的出生数，因此可以计算年龄别出生率。通过这一办法还可以为各人口亚群，例如按母亲的受教育程度，计算年龄别出生率。

421. 在亲生子女法中，任何年龄在15岁以下的人都可被称为“儿童”。之所以选择15岁这个年龄，是因为这大约是生育年龄范围的下限。对于每一个这样的儿童，首先要查明该儿童的母亲是否活着并住在同一住户内，其次，如果是这样的话，母亲是哪一个。住在家中的妇女的子女是这些妇女的亲生子女，而不是收养的或临时居住的其他妇女的子女；这一方法由此得名。“非亲生子女”是指生活在不包括其母亲的家庭的子女。

##### 1. 问题

422. 由于亲生子女法是存活倒推法的延伸，因此最基本的问题是那些有关出生日期或年龄的问题，这些问题已经在上文A.1节作了介绍。需要的其

他问题视情况而定，但可以包括“母亲的行号”、与户主的关系、出生婴儿数和尚存子女数。

423. 将子女与其母亲进行配对的最简单的方法是在普查或调查表上包括这样的问题：“此人的母亲的行号是什么？”“行号”指的是普查或调查表上包含有关母亲的信息的那一行的编号（联合国，1998年a，第2.75段）。表述“行号”可以采用一种行-列问卷格式，其中行对应人，列对应信息项目。其他格式可以通过适当地对这个问题进行重新表述来处理。

424. 如果有关母亲行号的问题是惟一的相关问题，那么可以告诉实地工作人员，如果孩子的母亲已经去世或住在别处，就让要填的空格空着。然而，在某些情况下，调查表将包括一个或多个有关父母是否活着的问题，以提供有关成人死亡率的水平和趋势的信息（见第七章C节）。在这种情况下，母亲行号的问题可以接着有关母亲是否活着的问题。例如，可以顺序地提出以下问题：

这个人的母亲还活着吗？

如果还活着，这个人的母亲生活在这个家庭吗？

如果是，这个人的母亲的行号是什么？

在这种情况下，由于一般会向所有人提出关于母亲是否活着的第一个问题，所以可以向所有人提出其他问题。获得的信息可以用于生育史重建，这是从普查和大规模人口调查中推出生育史的亲生子女法的延伸。生育史重建将在下面的章节介绍。

425. 提供有关成年妇女死亡率信息的关于母亲是否活着的问题通常伴随着一个关于父亲是否活着的问题，这个问题提供有关成年男子死亡率的信息。在这种情况下，在问及父亲时也可以像问及母亲时一样提出上面列出的后两个问题。这样就可以对父亲的特征以及母亲的特征计算出出生率，其中包括父亲的年龄别出生率，这是某些成年人死亡率估计方法所必需的，以及按父亲特征确定的人口亚群的常规的母亲年龄别出生率。

426. 即使有关母亲行号的问题不包括在普查或调查中，亲生子女法也可以常常使用。几乎所有人口普查和调查都包括一个有关与户主的关系的问题。这一信息常可用来推断子女与母亲之间的对应情况。

下文B节第4小节将进一步阐释的这一方法是令人满意的，条件是(a) 与户主的关系的问题提供足够的细节以及(b) 家庭比较接近核心（即，由母亲/妻子、父亲/丈夫以及一个或多个子女组成）。在成员人数较多的大家庭很普遍的情况下，如果使用亲生子女法，那么包括一个有关母亲行号的问题很重要。

表3. 亲生子女表

母亲 年龄	子女年龄													
	0	1	2	3	4	.	.	.	10	11	12	13	14	
15														
16														
17														
18														
19														
.														
.														
.														
60														
61														
62														
63														
64														
NOC														

注：NOC=非亲生子女

## 2. 制表

427. 一旦把子女与母亲进行了配对，就可以得到每一个亲生子女的母亲年龄，并且可以利用这一年龄来按子女年龄和母亲年龄为所有15岁以下的人编制以下表格。年龄必须按一岁分组。如果15岁以下妇女的生育数庞大，那么10-64岁妇女应当列入表中。50-64岁的妇女列入表中是因为这些妇女在普查前15年中生产了人口。“NOC”的意思是“非亲生子女”，即不与母亲一起生活的子女（不论因为母亲死亡还是因为她住在别处），因此对他们来说母亲的年龄无法确定。这一行事实上由母亲的年龄“不明”的那些子女组成。

## 3. 估计

428. 上一小节所述的表格用来计算普查前5年的年龄别出生率。在经过一些技术性调整后，表中的空格内的数字提供年龄别出生率的分子。对应的分母通过人口年龄分布的存活倒推获得。Cho、Retherford和Choe（1986年）的著作中详细介绍了这一方法。可以利用计算机程序来进行这些计算。还见联合国（1983年）。

#### 4. 配对

429. 本节提供的信息只有在查明家庭中母亲与子女的问题（例如“母亲的行号”的问题）不包括在普查或调查中的情况下才有关系。在这种情况下，可以利用人口普查和调查中几乎普遍包括的“与户主的关系”问题将母亲和子女在某些条件下进行配对。

430. 有关与户主的关系问题是一个向家庭中的每一个人提出的问题，询问他与户主或其他被调查人的关系。必须指定家庭中的一个人并且是惟一的一个人作为户主或被调查人。家庭中的其他人随后通过与此人的关系识别。为了避免不必要的重复，任何这类问题在这里将被称作是“与户主的关系”问题，即使问题指的是某种其他类别的被调查人。

431. 在与户主的关系的分类中可以发现许多变化，但所有变化都以亲属关系为主要依据。典型的类别有：户主、配偶、子女、子女的配偶、孙子女或曾孙子女、父母或配偶的父母、其他亲属、家佣及其他无关的人（联合国，1998年a，第2.73段）。

432. 下面各段落里简要概述了理想的配对情况。更多的解释见Cho、Retherford和Choe（1986）的著作第4章（1986年）。

433. 利用家庭中所有人的个人记录一个家庭一个家庭地进行配对。目标是确定，就当前家庭中每个15岁以下的人而言，是否(a) 这个人的母亲在家庭中以及，如果是的话，(b) 母亲是哪个人。对每个15岁以下的人依次得出答案。如果母亲不确定，那么这个人就被标明为“非亲生”子女。如果确定了母亲，则指明母亲的记录。

434. 例如，假设(a) 一个不满15岁的人与户主的关系类别是“子女”并且(b) 户主是一个15岁以上的女性。在这种情况下户主将被确定为孩子的母亲。

435. 或假设(a) 一个不满15岁的人与户主的关系类别是“子女”，(b) 户主是男性，以及(c) 家庭中有一名15岁以上的女性，与户主的关系类别是“配偶”。在这种情况下妇女可以被确定为15岁以下的人的母亲。当然，这名妇女有可能不是孩

子的母亲。她也许是第二任妻子，而不满15岁的人是第一个妻子的孩子。不过，在许多情况下，偶然的配对误差不会严重影响产生的估计数的准确性。

436. 可以利用各种补充信息来改善配对的质量。例如，可以核对母亲的年龄，以确保在用来同母亲配对的子女出生时母亲是处在生育年龄。如果可以得到有关尚存子女的信息，那么可以进行核对以确保没有妇女同超过其尚存子女数的数目进行配对。如果可以得到有关同母亲生活在同一个家庭中的子女的数目的信息（见下节有关出生婴儿数问题的讨论），那么可以进行核对以确保没有妇女与超过这个子女数的数目进行配对。

#### 5. 质量

437. 由于亲生子女法是存活倒推法的延伸，A节第4小节中对存活倒推估计的质量的探讨也适用于亲生子女估计以及存活倒推估计。更具体地说，亲生子女法得出的生育率水平估计与存活倒推估计容易出现的误差类型是一模一样的。

438. 亲生子女估计所表明的生育率年龄分布的质量反映子女与母亲配对的准确性以及母亲的年龄信息的准确性。年龄堆积的误差不会出现在按一岁分组的分年龄组出生率的估计中，因为特定年龄的母亲的更大数目一般会相应地与子女的大数目配对。然而，年龄的夸大会导致较大年龄组的年龄别出生率过高。为了用一个简单的人为编造的例子进行说明，如果所有40-45岁年龄组的妇女都被报告为45-49岁年龄组，那么对较大年龄组估计的生育率将大大高于真实生育率。

439. 如果使用有关母亲的行号的问题，那么一般有可能准确确定妇女作为家庭中每个孩子的母亲的被调查者身份。这个问题的目的是获取关于母亲-子女的生物关系的信息，但被调查者也许认为社会关系更为重要。例如，也许有一种趋势，把收养的子女报告为亲生的子女，即使在普查员作了相反说明之后仍然如此。如果，比如说，养母的年龄一般比亲生母亲大的话，这会扭曲生育率的年龄分布（Rindfuss和Sweet，1977年，第31页）。

440. 关于整个人口的生育率估计，子女与母亲配对的准确性只有在影响到被确定为每个孩子的

母亲的妇女的年龄的情况下才是重要的。如果不相干的妇女被确定为孩子的母亲，但她的年龄与真正母亲一样，那么估计不受影响。

441. 如果通过配对确定孩子的母亲，那么得出的估计数的准确性取决于对有关与户主的关系的问题的准确性以及这一信息及用来进行配对的其他信息的充分性。错误地报告与户主的关系不大可能是一个严重的问题。配偶、子女、孙子孙女和父母很少相互混淆。粗心大意的误差会导致这些类别中的某些人被确定为“其他亲属”，尽管这似乎不大可能。两类不相关的人之间的区别在某些情况下也许不清楚，但对于子女与母亲的配对并不重要。

442. 关于与户主的关系的信息以及使用的补充信息是否足以影响准确的配对，这主要取决于住户的构成。如果住户主要由核心家庭构成，那么配对常常能产生准确的结果。在住户包括祖父母、其他亲属和无关人员的情况下也可以得到准确的结果。然而，随着住户变得更加复杂，到了包括超过一个核心家庭的程度，那么母亲和子女准确配对的可能性就会降低。在这样的住户普遍存在的情况下，如果要采用亲生子女法，那么应该使用有关“母亲的行号”的问题。

443. 如果使用亲生子女法来提供对人口亚群的估计，那么误差可能产生于人口在这些亚群之间的移动。这一现象类似于在对国以下地理区域进行估计时移徙的影响，正如上文A.4节中联系存活倒推法进行探讨的影响一样。当人员在亚群之间转移时（例如改变就业情况或职业），对生育率的回溯估计将对那些普查或调查前各年不同的就业或职业情况群体中的人的经历进行混合。结果将减少群体间的估计差异。趋势以及水平可能会受到影响。

### C. 生育史重建

444. 生育史重建是亲生子女法的延伸，为来自人口普查或调查数据的65岁以下的每一个妇女得出完整的生育史（Luther和Cho，1988年）。重建的生育史可以用来得出许多像生育史一样详细的生育率估计数据。下文F节探讨了生育史。在这里，F节所叙述的通过详细提问得出的生育史可以被称作是“真实的”生育史。

445. 亲生子女法所涉及的将子女与母亲进行配对可以得出所谓的“亲生子女生育史”，这是年龄在15-64岁的每一个妇女的亲生子女名单，名单上显示每一个子女的性别和年龄。子女的年龄提供出生年份，而“年份”被理解是指开始和结束于普查或调查的参考时间的年份。

446. 如果对一些妇女来说，亲生子女的数目等于出生婴儿数，那么亲生子女史包括所有子女，因此是生育史。如果亲生子女的数目低于出生婴儿数，那么该妇女的一些子女居住在其他地方或者已经死亡。该妇女的亲生子女数与尚存子女数之差是她居住在其他地方的子女数。尚存子女数与出生婴儿数之差是她的死亡子女数。

447. 生育史重建根据亲生子女的出生年份和生育率及死亡率的估计年龄类型估计居住在其他地方的尚存子女和已死亡子女的出生年份。用这种方式可以得到每个妇女的完整生育史。对个别妇女来说，重建的生育史一般与她们的真实生育史不同，但重建的性质确保所有重建的生育史的总体情况在统计上与真实的生育史的总体情况相似。

448. 生育史重建的应用表明，当年龄报告比较准确时，从重建的生育史中得出的生育率估计值非常接近从报告的生育史中得出的同样的估计值（Luther、Feeney和Zhang，1990年）。

#### 1. 问题

449. 除了包括那些允许采用亲生子女法的问题外（见上文B节第1小节），普查或调查必须包括有关出生婴儿数（见下文D节）和尚存子女数的问题（见第六章A节）。

#### 2. 制表

450. 重建生育史的制表原则上与真实生育史的制表类似，但实践中有两个重要区别。第一个区别是重建生育史只提供每个子女的出生年份，而不是出生日期，其中年份指的是开始和结束于普查或调查的参考时间的年份。这在实践中给许多方面带来轻微的不便，但它需要针对这种情况设计的表格。

451. 重建生育史和真实生育史之间的第二个区别是，重建的生育史常常可以用于比真实的生育

史大得多的抽样规模。收集真实生育史时一般包括5 000到10 000个育龄妇女。而生育史重建可以用于任意大的人口调查和可能包括数百万个育龄妇女的普查数据样本。因此比起真实生育史来，从重建生育史中有可能产生更多、更详细的表格。

(a) 年龄胎次表

452. 重建生育史可以用来计算年龄胎次别出生率，其中“胎次”的定义是一个妇女生育的子女数。需要编制一系列的表格，每个胎次一张表。头胎见表4。由于根据重建生育史无法得出生育时母亲的年龄，所以母亲的年龄被计算到生育那年的年末。

453. 表4中的星号表示该表（可能的）非零空格。左上到右下对角线上的项目表示同年出生妇女组群的生育数。因此，对角线最底端的数据表示普查或调查时年龄为64岁的妇女的头胎；她们生头胎是在普查或调查前的第五十四到第一年。这意味

着这张表的编制与所示的完全一样，包括与10岁以下妇女的生育数对应的行，当然这些行包含的数据只能为零。这有利于根据该表计算年龄胎次别出生率所需的编程，并且还能检查年龄错报导致的误差。

454. 该表是关于妇女而不是出生人数。然而，由于在任何时期内生有某一胎次的妇女与同一时期该胎次的出生人数是一一对应关系，所以表中的数字可以被解释为头胎数以及生头胎的妇女人数。

455. 二胎表也一样（表5）。它也是有关妇女的表格，但空格表示二胎数以及生二胎的妇女人数。三胎或三胎以上的表与头胎和二胎表类似。

456. 系列表的最后一张表原则上将合并第n胎和第n胎以上的出生人数，之所以选择n是因为n胎和n胎以上的出生人数的比例小得可以忽略不计。实际上，最简单的办法一般是给头胎到第20胎制表。在大多数情况下，更高的胎次极其少见，完全可以不予考虑。

表4. 生有一个或多个子女的妇女，按头胎的出生年份和头胎出生年份末妇女的年龄列示

头胎的出生年份 (普查前……年)							头胎出生年份末 妇女的年龄		
54	53	52	.	.	.	3	2	1	
									0
									1
									1
									.
									.
									.
*	*	*	.	.	.	*	*	*	10
	*	*	.	.	.	*	*	*	11
		*	.	.	.	*	*	*	12
			.	.	.	.	.	.	.
			.	.	.	.	.	.	.
			.	.	.	.	.	.	.
			.	.	.	*	*	*	62
			.	.	.		*	*	63
			.	.	.			*	64

注：\* =可能填写的单元格  
 . =继续  
 空白 =没有登录



(b) 胎次表中的胎次-期间

457. 重建整生育史也可以用来编制按胎次和生育时母亲的“胎次期间”整年数列示的出生人数表。这些表可以用来计算Henry（1980年；最近的著作见Feeney，1983年；Feeney和Yu，1987年；Feeney，1988年；Luther、Feeney和Zhang，1990年；以及Feeney和Wang，1993年）发明的那种活产次数级数比率。

458. 活产次数级数一般指人口统计，该统计描述有一定数目的子女而又继续想要孩子的妇女比例，以及对于那些的确又有了一个孩子的妇女来说已定胎次与下一次生产之间的间隔时间分布。

459. 由于胎数为零的期间与年龄相同，所以到头胎的基础级数表与表4相同。从头胎到第二胎的基础表见图6。

460. 表的竖列，到生第二胎那年年末母亲拥有1胎的整年数，不过是将生第二胎的年份减去生第

一胎的年份。因此，生头胎和生第二胎发生在同一年的妇女（这一般是多胎的结果，只在极少情况下是生育间隔短的结果）在年末的胎次期间为零。在生头胎那年的第二年生第二胎的妇女在年末的胎次期间为1，依次类推。

461. 第二胎到第三胎的级数表与此类似（表7）。第三胎到第四胎的级数表以及更高胎次的级数表与头胎到第二胎的级数表和第二胎到第三胎的级数表一样。

462. 系列表的最后一张表原则上将合并从第n到n+1胎和所有更高胎次的级数的表。在实践中，最简单的办法一般是为一直到第19到第20胎的所有级数制表。更高胎次几乎可以完全忽略。

3. 估计

463. C节第3小节非常简单地叙述了对C节第2小节描述的表格的估计程序。应查阅所示参考书目以获得更多信息。

表5. 生有两个或两个以上子女的妇女，按第二胎的出生年份和  
第二胎出生年份末妇女的年龄列示

第二胎的出生年份 (普查前……年)									第二胎出生年份末 妇女的年龄
54	53	52	.	.	.	3	2	1	
									0
									1
									2
									.
									.
									.
*	*	*	.	.	.	*	*	*	10
	*	*	.	.	.	*	*	*	11
		*	.	.	.	*	*	*	12
			.	.	.	.	.	.	.
			.	.	.	.	.	.	.
						*	*	*	62
							*	*	63
								*	64

注：\* =可能填写的单元格  
 . =继续  
 空白 =没有登录

464. 由于这里描述的计算很少采用手工方法进行，所涉及的表的数目和规模尽管很大，但一般不成问题。

(a) 年龄胎次别出生率

465. 根据上文叙述的年龄胎次表（表4和表

5）计算的出生率是期间-同期群率，定义为用y年年初的妇女年龄x除这些妇女在y年期间的生育数。由于妇女的人数不随时间而变化，y年年初时的妇女年龄x的数目与y年期间这些妇女经历的人年数相同。

466. 中点（年龄-期间）比率习惯上被定义为y年期间年龄为x的妇女经历的人年数除y年期间生

表6. 生有两个或两个以上子女的妇女，按第二胎的出生年份和到第二胎出生那年年末妇女拥有1胎的整年数列示

第二胎的出生年份 (普查前……年)									第二胎出生那年年末妇女 拥有1胎的整年数
21	20	19	.	.	.	3	2	1	
*	*	*	.	.	.	*	*	*	0
	*	*	.	.	.	*	*	*	1
		*	.	.	.	*	*	*	2
			.	.	.	.	.	.	.
			.	.	.	.	.	.	.
			.	.	.	*	*	*	18
							*	*	19
								*	20

注：\* =可能填写的单元格  
 . =继续  
 空白 =没有登录

表7. 生有三个或三个以上子女的妇女，按第二胎的出生年份和第三胎出生那年年末妇女拥有2胎的整年数列示

第二胎的出生年份 (普查前……年)									第三胎出生那年年末妇女拥有 2胎的整年数
21	20	19	.	.	.	3	2	1	
*	*	*	.	.	.	*	*	*	0
	*	*	.	.	.	*	*	*	1
		*	.	.	.	*	*	*	2
			.	.	.	.	.	.	.
			.	.	.	.	.	.	.
			.	.	.	*	*	*	18
							*	*	19
								*	20

注：\* =可能填写的单元格  
 . =继续  
 空白 =没有登录

育时年龄为 $x$ 的妇女的生育数。 $Y$ 年中年龄为 $x$ 的妇女的中点比率约等于 $y$ 年年年初年龄为 $x-1$ 和 $x$ 的妇女的期间-同期群率的平均数。

467. 表4和表5提供(期间-同期群)年龄别出生率的分子。相应的分母是普查或调查时的各按一岁分组的妇女人数。由于表中的所有生育数都是对当时尚存的妇女而言,所以所有生育数的分母对任何同年出生的妇女都是一样的。因此普查或调查时年龄为64岁的妇女的人数是对应表中从右下到左上的对角线上的生育数的所有比率的分母。

#### (b) 年龄-胎次别出生概率

468. 用来估计年龄-胎次别出生率的表也可以用来估计年龄-胎次别出生“概率”。使用“概率”这个词是Whelpton(1954年)的发明并且是人口学所特有的。

469.  $y$ 年中年龄为 $x$ 的妇女的第 $i$ 胎的年龄-胎次别出生率与这一年这一年龄的年龄-胎次别出生概率之间的主要区别是比率的分母是所有妇女,而概率的分母只是能够生第 $i$ 胎的妇女,也就是说生有 $i-1$ 个子女的妇女。

470. 年龄-胎次别出生概率的分母的计算是通过从某年年年初经历相应产次的妇女人数中减去这一年中各胎次的出生人数,在同期出生群中循序渐进地进行的,这是Whelpton(1954年)发明的一种方法。例如,考虑某年年年初年龄为14岁的妇女组群,并且假定这些妇女中没有人生过孩子,那么所有人的胎数都为零。这些在年末时胎数仍为零的妇女的人数等于年初时胎数为零的人数减去(a)死亡人数和(b)在这一年中生育头胎的人数。在对普查或调查时尚存妇女的生育数采用这种计算方法时,死亡数不计,某年年末胎数为 $i$ 的妇女的人数通过从年初时胎数为 $i$ 的妇女人数中减去这一年中生育 $i+1$ 胎的妇女人数计算。

#### (c) 胎次和胎次期间别出生概率

471. 胎次和胎次期间别出生概率的计算原则上与年龄-胎数别概率的计算类似,但细节大不相同。年龄-胎次别概率的计算围绕妇女同期出生群,即同年出生的妇女群体进行。胎次和胎次期间别概率的计算围绕妇女的胎数分组,即在某年生有某

一胎次的妇女群体进行。Feeney(1988年)和Feeney及Wang(1993年)的著作中详细叙述了这些计算方法。

#### 4. 质量

472. 由于生育史重建是亲生子女法的延伸,所以,本章A节第1小节、第5小节和B节第5小节对质量的解释既适用于生育史重建的结果,也适用于存活倒推法和亲生子女法的结果。除了这些小节中探讨的考虑事项,从重建生育史中得出的估计数的质量主要取决于非亲生子女和死亡子女出生日期估算的质量。当死亡率为中低水平并且大多数15岁以下的子女与其母亲生活在同一个家庭中时,只需对所有出生人口的一个较小部分估算出生年份,这样一来即使不完善的估算程序也会得出比较可靠的总体结果。Luther、Feeney和Zhang(1990年)的研究表明,根据重建生育史得出的估计数几乎可以像根据真实生育史得出的估计数一样准确,条件是存活倒推法和亲生子女法得出可靠的结果。

#### D. 平均生育数

473. 有关妇女的平均生育数的问题在过去50年的人口普查中广泛使用,并且总的说来非常成功。这些问题还被成功地用在人口调查中。它们提供有关生育率的重要信息,在与有关尚存子女的问题相结合(第六章,A节)时,还能提供有关婴幼儿死亡率的重要信息。

474. 有关平均生育数的数据对于那些没有完整的出生登记的国家很重要,在这些国家,这些数据成为有关生育率水平和趋势的替代信息来源。不过,这些数据对于有完整的出生登记的国家也很重要,原因有两个。第一个原因是根据妇女胎次和胎数计算生育率估计值一般需要这些数据,这样的计算对于分析生育率趋势也许非常重要。有关出生的民事登记数据可以很容易地按胎次及年龄制表,但相应的“面临风险”妇女的数目通过按年龄和胎数分列的妇女的普查表提供。第二个原因是来自普查或调查的平均生育数数据可以用来参考在普查或调查中收集的其他信息研究差异生育率。特别是,一般说来,社会经济信息在普查和调查数据中远比在生命登记数据中详细。

## 1. 问题

475. 基本的问题是：

### 平均生育数

你[这名妇女]一生中总共有过几个子女？包括出生后不久夭折的子女和居住在其他地方的子女。

如果可能的话，应该向所有也许生过孩子的妇女问这个问题。在某些情况下，只向结过婚的妇女问这个问题；在其他情况下，只向超过某一最低年龄的妇女问这个问题，最低年龄一般定在10到15岁之间。

476. 在少数情况下，规定了一个年龄上限。建议不要这么做：年龄上限会诱使实地调查夸大妇女的年龄以避免问这个问题；另外，老年妇女的信息常常提供从其他来源得不到的有关生育率的历史水平和类型的信息。

477. 有关平均生育数的问题常常与第六章A节进一步探讨的有关尚存子女数的问题一起提问。在这种情况下，两个问题可以连在一起提问：

### 平均生育数

你[这名妇女]一生中总共有过几个子女？包括出生后不久夭折的子女和居住在其他地方的子女。

### 尚存子女

这些子女中有几个现在还活着？

这两个问题有时发展为以下四个问题：

### 居住在这里的子女

你[这名妇女]生育的子女中有几个住在这个家里？

### 住在其他地方的子女

你[这名妇女]生育的子女中有几个住在其他地方的子女？

### 死亡的子女

你[这名妇女]生育的子女中有几个已经不在人世？包括所有活产子女，其中包括任何出生后不久即夭折的子女。

### 平均生育数

在你[这名妇女]的一生中总共生育过几个孩子？

从总体上看，这四个问题在逻辑上显得罗嗦，因

为对任何三个问题的回答都必然包含对第四个问题的回答。不过，有必要清楚地表明平均生育数包括已经死亡的子女，并且重复可以提供检查，在实地工作中用来查明与调查前后不一致的回答。在某些情况下，这张清单上还要加上有关死产婴儿数的问题，例如在1970年马来西亚的普查中就加上了这样的问题（Cho, 1976年）。

478. 在人口调查中，如果有可能包含更多的问题的话，最好是在详细询问出生婴儿数之前问一个“过滤性”问题，例如下面这个问题：

你[这名妇女]有过活产经历吗？

这个问题可以包括“活产”的定义（见术语表中关于活产的条目）。

479. 问四个问题而不是两个问题的根本原因是，增加的问题可以通过鼓励加上死亡子女和生活在其他地方的子女来提高平均生育数报告的完整性。四个问题是不是真能得到足够优良的结果来证明增加两个问题的负担是有理由的，这一点尚未得到以经验为依据的研究的证实。在包含较多问题的人口调查中，也许用四个问题是可取的。在通常需要严格限制问题总数的人口普查中，选择不那么明显。在没有实地试验结果来清楚地表明四个问题得出的结果更好的情况下，也许只用两个问题更好。

480. 有关平均生育数的问题可以进行扩展，以得到有关男孩和女孩的人数的信息，这样就可以估计出生时性别比率，如果还查明了尚存的男孩和女孩的人数的话，还可以估计死亡率的性别差异。不过，出生时性别比率在人口间的差异不大，因此为查明出生婴儿的性别而提出的问题只有在报告的质量比较可靠的情况下才有用。这一普遍规则的一个重要的例外情况发生在存在流产的性别差异的情况下，这可以导致出生时性别比率极高。

## 2. 制表

481. 尽管半个多世纪以来，全世界的普查和调查中都包括有关平均生育数的问题，但有关平均生育数的数据表涉及好些微妙之处，令人遗憾的是，公布的令人不满意的表的例子如出一辙。问题包括未定后限的年龄组和胎数组的选择、对不确定情况的处理以及对妇女生育的子女总数的规定。

482. 表8可以被看作是对出生婴儿数数据制表的一般性指导。

制表的具体规定取决于被问及平均生育数问题的妇女。如果向超过某一年龄的所有妇女提这个问题，那么标题将提及“……岁及以上妇女”并且年龄分类的底限将反映这一点。如果只向结过婚的妇女提出这个问题，那么标题将提及“结过婚的妇女”并且年龄分组将反映被问及婚姻状况的妇女的年龄底限。

483. 标准年龄分组是按五岁分组，必要时可以进行修改以反映被调查妇女的年龄。未定后限的年龄组应当定得高一些，而不是低一些，以免不必要地丢弃可能证明是有用的数据。建议把最低85岁作为未定后限的年龄组的下限。对于85岁以上人群比例较大的人口来说提高年龄也许是适当的。如果不确定年龄经过了估算的话，不确定年龄（NS）一行将不存在。

484. 明智的做法是至少按一岁分组为全国范围内的平均生育数编制一张表。如果年龄报告准确，那么该表将提供比只显示五岁年龄组的表更多的信息。如果年龄报告质量差，那么按一岁分组的表有助于评估年龄错报对按母亲年龄列示的平均生育数分布的影响（Hull和Sunaryo, 1978年；Hull, 1995年）。

485. 平均生育数应该以单一的整数表示，决不要分组，除非是对于未定后限的年龄组群。在无限限制胎数组的情况下，一般舍去胎次分布，因为变通办法是允许尽可能大的平均生育数，这个数大于30（Bongaarts和Potter, 1983年）。对未定后限的年龄组的选择应使这一组中的妇女的比例低于所有年龄组的1%。对于高生育率的人口，无限限制胎数组一般应不早于15个或更多的子女。

486. 如果对不确定的平均生育数的情况进行了估算的话，“不确定平均生育数”一栏将不存在。如果不确定的情况存在，那么必须明确地显示出来。不确定的情况决不应与零胎数妇女合为一组。在这样做的一个有文献记录的例子中，没有子女的妇女的比例被大大高估了（Feeney, 1991年）。

487. 计算任何特定年龄组的妇女的平均生育总数可以根据这些妇女的未截尾分布，用生有一个子女的妇女的人数加上生有两个子女的妇女人数的两倍，加上生有三个子女的妇女人数的三倍，依次类推。完整的分布对于准确的计算来说是必要的，因为对于截尾分布，没有办法知道在无限限制胎数组中的妇女的平均生育数。

表8. 按年龄和平均生育数列示的妇女人数以及按妇女年龄组列示的平均生育总数

妇女年龄	平均生育数								不确定平均生育数	妇女总数	平均生育总数
	0	1	2	.	.	.	14	15+			
10-14											
15-19											
20-24											
.											
.											
.											
80-84											
85+											
NS											
总计											

注：CEB=平均生育数  
NS=不确定

488. 显示平均生育总数的栏目应被看作是强制性的，因为遗漏这一栏常常会影响从表中得出的估计数的质量，并且一定会大大增加计算估计数的工作。这一栏中的总数是在制表期间得出的，方法是：在每次处理记录时把平均生育数加入到适当的单元格中。

### 3. 估计

489. 有关平均生育数的数据为估计各种生育率统计数据提供了基础。只有生育史数据，不论是真实的还是重建的，为得出详细的生育率统计数据提供更多的机会。不过，可以通过单独的一个问题得出有关平均生育数的数据，而生育史数据不是需要询问许多复杂的问题（下文F节），就是需要采用一种复杂的方法（上文C节）。D节第3小节简要叙述了可以根据平均生育数数据估计的各种生育率统计数据。

#### (a) 胎次分布

490. 可以从平均生育数数据中估计出的最简单的生育率测度是各年龄组中妇女的胎次分布。在这里“胎次”是“出生婴儿数”的同义词。胎次分布不过是任何年龄组中的妇女按平均生育数显示的比例分布。

491. 如果没有“不确定的”情况，那么胎次分布就是用妇女总数除有各个生育数的妇女的数目。如果有“不确定的情况”，那么应该用妇女总数减去生育数不确定的妇女的数目，然后再除。

492. 然而，对于后一种计算要特别当心。有可能出现这样的情况，即与生过一个或多个子女的妇女相比，没有孩子的妇女未报告生育数的情况要多得多。如果是这种情况，那么应试图估计出没有说明生育数的无子女妇女的比例。详情见下文第4(b)小节。

493. 胎次分布对还没有超过生育年龄的妇女来说意义不大。不过，对于接近或过了生育年龄的妇女来说，胎次分布按育龄末生育率估计妇女的分布情况。这样的育龄末生育率分布提供有关生育率高和生育率低的人口生育率类型的重要信息。对于生育率高的人口，育龄末生育率显示相当一部

分妇女只有少数几个孩子，尽管一般说来妇女也许生有六个或六个以上的子女。就生育率低的人口而言，育龄末生育率显示人口中三胎和三胎以上的比例高，其中育龄末平均生育率低于每个妇女两个孩子。好几卷《联合国人口年鉴》都具体提到了生育率，涉及过去半个世纪中全世界几十个国家，其中的平均生育数数据证明了前面的结论。

494. 可以对能够得到平均生育数的任何妇女群体计算胎次分布，包括所有妇女和各种已婚期间组群中的妇女以及各种年龄组中的妇女。不过不按年龄或已婚期间划分的妇女组群的胎次分布一般意义不大，因为刚开始生育的、没有子女或子女不多的妇女与育龄末妇女的经历被混为一谈。

#### (b) 活产次数级数比率

495. 活产次数级数比率是至少有特定数目的子女的妇女中继续再生一个孩子的妇女的比例；或者，就没有子女的妇女而言，至少生一个孩子的妇女的比率。

496. 在根据表8所示的平均生育数表格计算活产次数级数比率时，从第*i*胎到第*i+1*胎的级数比率是用生有*i*胎或以上子女的妇女人数除生有*i+1*胎或以上的妇女人数。如果*i=0*，那么这一比率与用妇女总人数除有一个或一个以上子女的妇女人数得出的比率相同。

497. 可以对能够得到平均生育数的任何妇女组计算活产次数级数比率，但对于过了生育年龄或就要过生育年龄的妇女来说最有意义。在实践中，这意味着妇女接近或超过50岁。

#### (c) 平均生育数

498. 各年龄组妇女的平均生育数便于对人口生育率的总结。这些平均数的计算是用某个年龄组中的妇女总人数除这个年龄组中妇女生育的子女总数，同时适当考虑不确定生育数的妇女。

499. 如果必须根据生育数的截尾分布计算生育总数，那么可以利用各种办法来大致估计无限胎数组中妇女生育的子女数。一种方法是计算和推断活产次数级数比率。一种更简单的方法是不那么现实地假定较高胎次的活产次数级数比率完全相等，

1947年和1957年马来亚人口普查的研究报告中叙述了这种方法（Saw Swee-Hock, 1964年）。只有在无限组中妇女比例小的情况下，这些方法才能对出生婴儿总数作出令人满意的估计。避免这种估计方法是把“出生婴儿总数”一栏列入出生婴儿数表格的一个重要理由。

#### (d) 完全生育率

500. 完全生育率是指达到育龄末（一般定为50岁）的妇女的真实组或假定组的平均生育数。例如，在一次特定的普查中年龄在50-54岁的妇女的平均生育数是这组妇女完全生育率的估计。

501. 完全生育率也可以指同期出生群，也就是说在某个时期出生的妇女群体。这是同期出生群和年龄组之间的简单对应，因为在任何特定时间任何年龄组中的人都出生在一个特定的时期，因此属于一个特定的同期出生群。例如，2000年1月1日年龄在50-54岁的妇女一定是在日历年1945-1949年之间出生的。

502. 然而，这一同期出生群的完全生育率不一定是这一年龄组的完全生育率，因为没到50岁就死亡的妇女的平均生育数也许与活到50岁的妇女的平均生育数并不相同。

503. 完全生育率有时通过累计一个妇女同期组群的分年龄组出生率计算。在这种情况下，完全生育率指的是一个有这些分年龄组出生率但没有任何死亡率的假定的妇女组群的平均生育数。

504. 50岁以上各年龄组的妇女的平均生育数常常被看作是对相应的同期出生群的完全生育率的估计。在生育率和死亡率水平不太高的情况下，这些估计数的质量可能很好。

505. 50岁以下的各年龄组的妇女的平均生育数也可用来估计相应的同期出生群的完全生育率。例如，考虑在某次普查或调查时年龄在25-29岁的妇女组。这些妇女还没有过生育年龄，因为将来她们还会生孩子。然而，假设可以得到该同期组群的生育率年龄类型估计。根据该年龄类型，有可能计算这一同期组群中的妇女在25-29岁时的平均生育数与该同期组群的完全生育率的比率。《手册十：人口间接估计技术》（联合国，1983年）中叙述的P/F比

率法使用的主要就是这一计算方法。用这个比率除年龄组中妇女的平均生育数就得到相应的同期出生群的完全生育率的估计数。进行这一估算所需的生育率年龄类型可以通过不同的方式得到。最简单的方法是利用普查或调查前一年根据亲生子女法或有关最近的生育数的问题得到的生育率的年龄类型（见下文E节）。可使用更复杂的方法考虑生育率年龄类型随时间发生的变化。

#### (e) 总和生育率

506. 任何同期出生群的完全生育率提供直至该同期出生群达到平均生育年龄时的总和生育率估计（有关总和生育率和平均生育年龄的定义见术语表有关条目）。这一观察为根据有关出生婴儿数的普查或调查数据估计生育率趋势提供基础。由于平均生育年龄（一般略低于30岁）即使在生育率发生变化时也常常变化不大，因此估计的质量主要反映平均生育数的准确性。在可以利用来自两次或多次普查和/或调查的有关出生婴儿数的数据时，这一方法特别有用，在这种情况下，比较各次普查或调查得到的趋势提供了一种评估估计准确性的简单方法。

507. 总和生育率也可以通过下面段落中说明的方法来估计。

#### (f) 年龄别出生率

508. 当可以利用从连续两次或多次普查和/或调查中得到的按母亲年龄分列的有关出生婴儿数的数据时，可以区分同期出生群的平均生育数，以估计该同期组群在普查和/或调查间隔期间的年龄别出生率。可以使用各种技术从该数据中得出期间生育率估计。例如，见Coale、Cho和Goldman，1980年。

## 4. 质量

#### (a) 出生婴儿漏报

509. 1950年代和1960年代非洲的经验显示(a) 尽管近期出生人数常常严重漏报，但生育率年龄类型可能还算准确；并且(b) 对于年轻妇女来说，出生婴儿数的报告的完整性常常令人满意，但

老年妇女的情况就差多了。这导致了P/F比率法的发明，用来在可以利用两种类型的数据时估计生育率水平（Brass等人，1968年；联合国，1983年）。这种方法利用来自最近出生人数的数据生育率年龄类型信息和来自出生婴儿数数据的生育率水平信息，同时假定生育率水平没有变化，来估计和纠正两种误差。

510. 早期经历的老年妇女的生育数漏报导致不恰当地忽视了这些数据，以后的经历显示这些数据也许像有关年轻妇女的数据一样准确。例如，一项研究使用1962、1969和1979年肯尼亚人口普查中50岁以上妇女的出生婴儿数数据来显示肯尼亚的生育率水平在1920年和1960年之间大幅度上升。尽管来自所有三次普查的数据都远远不够完美，但误差与随着年龄增长报告的完整性下降模式不相符，并且尽管有误差也可能得出有关生育率水平和趋势的可靠结论。鉴于1960年以前该国的数据缺乏，以任何其他方式大概无法得出这一结果。

511. 如果被调查者未能报告已死亡子女和生活在其他地方的子女，因为他们不清楚这些子女应该包括在内，那么上面讨论的更详细的问题可以改善回答的质量。不过，用较简单的问题得到了可靠的结果，而用较复杂的问题得到的结果质量较差。在问题的含义上混淆不清不是漏报的惟一解释。例如，被调查人也许不愿意提起已经死亡的子女，在这种情况下阐明问题的含义不可能改进回答。

512. 被调查者在报告出生婴儿数时也许把收养的子女包括在内，而把送给别人抚养的亲生子女排除在外。这会导致个别妇女的报告出现误差，并且会扭曲胎次分布的情况（如果无子女是收养动机的话）。不过，这不会影响出生婴儿总数的报告，除非生活在别的国家的妇女生育的子女数量庞大。不建议增加有关被收养的子女的具体问题。在这么做的一个例子中，报告的质量如此糟糕以致得出的数据无法利用，尽管在以前的普查中成功地使用了出生婴儿数和尚存子女数的基本问题（Banister，1979年）。

#### (b) 点查员未能记录无子女的情况

513. 应训练实地工作者对没有子女的妇女清楚地记录出生婴儿数为“0”。也许会以为这是显而易见的，但经验显示实地工作者在遇到没有子女的妇女，特别是非常年轻的妇女时，常常把出生婴儿数一栏空着不填，或者填“-”，而不是“0”。因此，在处理数据时，实地工作者正确查明或本来能够正确查明实情的大量没有子女的妇女被不正确地划入“出生婴儿数不明”一类。

514. 这一误差导致无子女妇女的比例过低，平均生育数过高。平均生育数中的误差大不可能很严重，除非对非常年轻的妇女而言，但无子女妇女的比例中的误差也许会严重到使这一信息无价值可言。

515. 对这一现象的认识至少可以追溯到50年以前——1947年马来亚普查主要报告清楚和详细地描述了这一现象（Del Tufo，1949年）——但导致的误差在许多普查中都可以看到。Vincent（1946年）首创了发现并且在某些情况下纠正列表数据中的这些错误的技术，El Badry（1961年）详细阐述了这一技术。

(c) 选择误差

#### (c) 选择误差

516. 然而，报告准确无误的出生婴儿数数据只代表普查或调查中包含的妇女的经历，可一般情况下对这些数据的处理就好像它们代表一般人口过去的经历一样，包括在进行普查或调查前已经死亡的妇女。一般说来，这不会引起问题。就高龄妇女而言并且在非同寻常的死亡模式的情况下会出现例外——例如，就获得得性免疫缺陷综合症(艾滋病)引起高死亡率的人口而言。

#### E. 最近出生人数

517. 可以询问普查和调查对象在过去一年或其他最近期间其家里或家里的某个妇女有多少活产人数。可以用其他问题来查明是什么时候生的，出生婴儿的性别以及婴儿是否存活。或者，对于至少有一个子女的妇女，可以问被调查者最近一次生育的日期。所有这些问题都是为了获得有关“最近出生人数”的信息，与有关出生婴儿数的问题不同，这些问题需要被调查者回忆某次生育发生的时间。

518. 有关最近出生数的问题所提供的数据也可以通过运用亲生子女法获得（上文B节），这一



方法不仅将提供有关最近出生人数的信息，而且还将提供有关普查或调查前15年出生人数的信息。因此，如果运用亲生子女法，那么在普查或调查中包含有关最近出生数的问题也许没有什么价值。

## 1. 问题

519. 第一个和最简单的问题是：

- (1) 你[这名妇女]在过去12个月有过活产吗（如果是多胎产，请说明人数）？

520. 向所有已婚妇女或所有适龄妇女提出这个问题。该问题的字面意思是“在过去12个月你有过几个活产？”但生过不止一胎的妇女人数极少，因此这么问有可能会让实地工作人员和被调查者都觉得奇怪。原则上应允许多胎产的可能性。不过，多胎产的机率一般非常低，这一细节在实践中常常被忽视。可以对这个问题进行扩展，问问出生子女的性别或是否还活着。

521. 将得到类似信息的第二个问题是：

- (2) 你最近一次活产是在哪年哪月（如果是多胎产，请说明）？

这个问题应该只问生有一个或多个子女的妇女，因此只使用有关出生婴儿数的问题才合适。多胎产的可能性原则上应该允许，但这一限制在实践中似乎常常被忽视。该问题不会产生普查或调查前一年有关出生人数的有用信息。早年报告的出生人数不完整，因为一些出生婴儿的母亲会再生孩子并将报告这些孩子的死亡年月。这个问题也可以扩展，询问孩子的性别以及是否仍然活着。

522. 将得到类似信息的第三个问题是：

- (3) 在过去12个月这个家庭中有过任何活产吗？

这个问题指的是在过去12个月的任何时间家庭成员生育的子女，不管他们是否还活着或是否仍是家庭的成员。接下来必须询问报告的每次活产时母亲的年龄。经常也会询问有关报告的每个出生婴儿的性别的问题。

523. 原则上该问题将提供有关后来死亡或离开家庭的妇女在家庭里生育的婴儿的信息。不过，

由于观察单位是家庭中发生的活产而不是生活在家庭里的妇女（就像在第一个问题中一样），调查问卷需要有一个独立的部分，并且必须包括有关生育时母亲年龄的补充问题。这类有关最近生育数的问题在形式上与下一章探讨的有关最近死亡数的问题类似。

524. 第一个问题和第三个问题可以询问除普查或调查前12个月以外的某一时期内的生育情况，例如，普查前最近的完整日历年或者普查前18个月或24个月。在这种情况下，获得有关生育时间的进一步信息也许有用，例如生育发生在哪个季度或半年。

## 2. 制表

525. 在询问妇女过去12个月（或其他期间）生育了多少子女时，基础表格显示按生育数和普查时妇女年龄分列的被问及该问题的所有妇女。如果获得一个较长期间及/或半年或季度的信息，表格应当扩展，以显示有关调查问卷的可获得的全部详细信息。如果得到出生婴儿的性别和/或存活情况的信息，出生婴儿的男女人数和/或存活人数及死亡人数应当显示在表格中。

526. 应对全体人口编制按一岁分组的表格。如果年龄报告的质量可靠，那么该表将包含有关生育率年龄类型的有用信息，如果只按五岁分组列示数据，那么这一信息会变得费解。如果年龄报告质量差，那么按一岁分组的数据将有助于评估其影响（Hull和Sunaryo; Hull, 1995年）。

527. 在使用有关最后一次生育的月和年的问题时，基础表格显示按普查或调查时妇女年龄以及最近一次生育的年份/时期分列的生有一个或多个子女的全体妇女。一些年月组是恰当的，因为使用的主要是普查或调查前一年的出生婴儿数。同样应该为全体人口编制按一岁分组的表格，但按五岁分组的表格对于国以下地理区域以及人口亚群来说也许就足够了。如果这个问题与有关出生婴儿数的问题一起使用，正如通常情况下所做的那样，还应该编制按年龄、最后一次生育的年份以及出生婴儿数列示的生有一个或多个子女的妇女的表格（Feeney和Ross, 1984年）。

528. 根据有关家庭生育数的信息编制的表格在形式上与根据生活在家庭内的妇女的生育数问题编制的表格类似。由于信息来自调查问卷的一个独立的部分，它通常将包括在单独一套数据中，这套数据含有有关报告的每个出生婴儿的一个记录。因此表格的单位是出生婴儿而不是妇女。表格将按生育时母亲的年龄分类显示普查或调查前12个月的出生婴儿数。如果得到有关出生婴儿数的其他信息，可以为这一信息加上其他栏目。

### 3. 估计

529. 经验显示有关最近出生婴儿数的问题有可能导致出生婴儿数漏报，尽管细致的实地调查可以得到90%或更高的完整率。因此，有必要借助某些手段来评估报告的完整性程度，以便上调报告的数字。为此发明了P/F比率法。在其最初的形式中，这一方法要求普查或调查前几十年的生育率恒定不变。后来的发展放松了这一假定（联合国，1983年，第二章，B节）。

530. 在使用有关妇女生育数的问题而不是家庭生育数的问题时，表格的年龄范围指的是普查或调查时妇女/母亲的年龄，而不是她在生育时的年龄。因此，对年龄在15-19岁的妇女计算的出生率一般指从14 1/2开始到19 1/2结束的年龄段，对于较大年龄组来说也一样。可以用不同的方法转换为标准的按五岁分组的年龄组，最常见的是利用生育率年龄类型模型。

### 4. 质量

#### (a) 报告误差

531. 普遍趋势是漏报出生人数。鉴于生育经历的性质及其对家庭和妇女及其家人生活的重要性，很难想象会在这么短的时间里忘记生育情况。漏报一定另有原因。

532. 一种解释是实地调查工作者表现不佳，如果回答是否定的，那么他们就能省点事。公然忽视是一种可能，但没有经过适当训练的实地调查工作者也许以为通过“预测”被调查者的情况可以提高收集数据的效率。他们也许认为，对于年龄较小或较老的妇女或者身边没有幼儿的妇女，他们知道

问题的答案，因此没有必要提问。在对实地调查工作者进行培训和监督时应涉及这些问题。

533. 漏报出生婴儿数的另一种解释是被调查者记不清生育日期，因此对生育是否发生在问题所指的期间没有把握。不过，这种“参考期间误差”既容易导致多报也容易导致漏报出生婴儿数，而多报的情况很罕见。一种可能的解释是不确定导致被调查者回避这个问题，于是实地调查工作者记录无出生婴儿。

534. 询问最近生育的年和月的一个根本原因是，这个问题要求每个合乎条件的妇女正面回答这个问题并由实地调查工作者清楚地登记在普查表上。这可以防止被调查者或实地调查工作者企图回避生育事实的情况。不过，无答案始终是一种选择，不管是被调查者还是不问这个问题的实地调查工作者都可以作这样的选择，因此这一方法不一定能产生可靠的结果。

535. 漏报出生婴儿数的第三个解释是，被调查者也许不愿意报告已经死去的出生婴儿。在这种情况下，被调查者也许报告存活的最小子女的出生年月而不是最后一次生育的年和月。由于这些出生婴儿即使不是全部也是大多数出生在早先时候，这将导致普查或调查前12个月期间的出生婴儿数漏报。询问最后一个出生的子女是否存活的一个原因是检验这一形式的回答误差，存活的子女比例过高将表明存在这样的误差。

536. 在使用有关家庭中出生婴儿数的问题时，对报告误差的考虑一般与在使用有关妇女的生育数问题时类似。不过，在把有关家庭中出生婴儿数的问题与有关家庭中死亡婴儿数的问题结合起来提问时，也许会问被调查者他们是否确定没有未存活下来的活产婴儿。随后可以使用这一调查结果来修正有关最近出生婴儿数和最近死亡婴儿数的信息（见第六章C节）。

537. 利用有关最近出生婴儿的性别分布的数据可以计算出出生时性别比率。不过，由于出生时性别比率的变化范围一般较小，计算得出的出生时性别比率就绪的最大用处也许是表明男女出生婴儿的报告完整性差异。然而，在女胎流产比例高的地方，可以发现非常高的性别比率。

## (b) 选择误差

538. 有关妇女在过去12个月（或其他期间）的活产数的问题以及有关她们最近一次活产的年和月的问题显然排除了在普查或调查前死亡的妇女的生育数。有关家庭中发生的活产数的问题（如果回答准确的话）能得到这些妇女的生育数。不过，家庭问题不会得到在普查或调查前解体的家庭的生育数，这些生育数将（如果回答准确）通过有关妇女健在时出生婴儿数的问题得到。在实践中这些问题的区别也许不太重要，因为有关最近生育数数据中的选择误差有可能被报告误差掩盖。不过，在特殊情况下，选择误差可能变得非常重要，例如在人体免疫缺陷病毒/获得性免疫缺陷综合症(艾滋病病毒/艾滋病)广泛流行的情况下。

## F. 生育史

539. 生育史主要包括一名妇女所生的全部仍然活着或已经死亡的子女名单以及有关这些子女的某些信息，信息包括他们的出生日期、性别以及是否是多胎产中的一个。虽然叫“生育史”，但生育史是有关婴儿和儿童死亡率的一个重要的信息来源（第六章）。

540. 生育史绝对是人口数据收集活动中使用的最深入细致的提问形式。其使用一般限于5 000到10 000个妇女样本，尽管偶尔也在较大范围的调查，有时甚至在人口普查中使用。例如，1982年中国进行的调查收集了超过300 000名妇女的生育史（中国人口信息中心，1984年）。

541. 有关生育史的大多数经验不是在世界生育率调查方案中就是在人口和健康调查方案中。

### 1. 问题

542. 包括生育史是生育率调查的一个限定特点。生育率调查一般包括两个基本问卷：一个是相对简单的家庭问卷，类似于在较大范围的人口调查中使用的那些问卷；另一个是详尽得多的妇女问卷，包括生育史的问题。妇女问卷常常针对家庭问卷上出现的年龄在15-49岁的已婚妇女。有时向50岁以上的妇女收集生育史信息，特别是在中国1982年的千分之一人口调查中（中国人口信息中心，

1984年），但这很罕见。

543. 生育史问题几乎总是接着上文D节第1小节叙述的有关出生婴儿数和尚存子女数的详尽问题提出的。按照这种方式，在详细的生育史问题提出之前先查清了妇女的出生婴儿总数，对于这些妇女应获得生育日期和其他信息。

544. 随后对每一次活产提出一系列问题，其中包括如下：

孩子姓名

是男孩还是女孩？

孩子出生在哪个年月？

孩子是双胞胎中的一个吗？

孩子还活着吗？

如果已经死亡，他/她当时多大？

出生日期（例如出生年月日）可以用于特殊的目的。如果可以获得有关出生日期的可靠报告，那么最好是询问出生日期（年月日）而不只是年和月。

545. 可以指示调查员交叉核对活着的子女与家庭问卷上记录的子女，并且在不一致的情况下要求被调查者说明原因。原因通常是子女生活在其他地方。

546. 如果孩子出生在出生后一个月内夭折，那么死亡子女的年龄应以天数报告，如果子女出生后两年内死亡，那么应以月数报告，如果子女出生两年后死亡，那么应以年数报告。获取两年而不是一年内死去的子女的死亡月份的主要原因是，如果对于刚刚超过以及差一点不到一年的年龄可以获得以月表示的死亡年龄的话，那么对常见的大量以月表示的死亡年龄的评估会更加可靠。尤其是，将死亡时的年龄归算到12个月，就会将大量不足一岁死亡的人数归入儿童死亡这一类别，从而减少了婴儿死亡人数。

547. 在某些情况下，可以使用有关死亡年月的问题来替代有关死亡年龄的问题，不过，其弊端是对于在出生后一个月内死亡的子女不那么容易获得以天数表示的死亡年龄数据。

548. 生育史一般包括一直到调查时妇女生育的所有子女，但在某些情况下，生育史被局限在过去5年或10年中生育的子女。局限于部分生育史的主

要理由是减轻收集有关全部子女信息的负担，同时得到大多数好处。部分生育史的风险是调查员会写错子女的死亡年龄或日期来避免收集生育史所要求的更详细的信息。一般认为节约抵不上风险，因为收集全面的而不是部分的生育史所增加的成本一般很小。

549. 也可以通过收集信息时仅限于胎次来获得部分的生育史。例如，部分生育史可以包括有关从头胎到第四胎以及有关最后一胎的信息，但不包括除最后一胎外的第五胎及以后胎次的信息。这种方法不适合高生育率的国家，但对低生育率的国家也许适合。

550. 有时使用包含妇女在她一生中所经历的所有妊娠的信息的妊娠史。在许多情况下，对没有导致活产或死产的妊娠的报告非常不全面。通过确保不混淆死产与婴儿死亡，甚至报告不全面的死产也可以用来得到更准确的婴儿死亡率，并且足够详尽的实地程序在某些情况下可以确保足够全面的死产记录。

## 2. 制表

551. 对生育史数据的处理方式一般不同于来自上文探讨的其他问题的数据的处理。对于前面各节探讨的各个问题，个别普查或调查记录被用来编制表格。这些表格随后用来——常常同也许来自其他来源的其他表格一起——得出生育率和/或死亡率水平和趋势的估计数。

552. 相反，一般对生育史数据进行处理，以直接得出生育率和死亡率的合理的估计值。可以通过进行数据处理的计算机程序编制表格，作为中间结果，但一般不输出这些表格。一个重要的例外情况发生在用来查明某几种回答误差的特殊表格方面。例如，参见国际人口学研究所公布的报告中的数据质量表（Rethornd, Mishna和Prakasam, 2000年，附录D）。

## 3. 估计

553. 根据生育史数据计算分年龄组出生率和婴儿及儿童死亡率原则上很简单，但实际做起来很复杂。之所以很复杂在很大程度上是因为只有有关出生年月的信息而不是有关出生日期的信息。有关的复杂之处在于计算生活的人年数，这是各种比率

的分母。计算原理在Verma（1980年）的著作中有详细叙述。

## 4. 质量

### (a) 报告误差

554. 生育史数据中的最严重的报告误差涉及出生日期的错报，这导致大量出生人数从调查之前年份转移到较早的年份。结果是低估调查之前年份的生育率水平和高估较早年份的水平。结果将造成生育率下降的表象，而事实上生育率保持稳定（或者甚至在上升）或者夸大实际的下降幅度。

555. 误差可以是大规模的并且是致命的，因为观察家们也许渴望生育率下降的情况发生并且会不加批评地接受这一现象发生的证据。结果，下降的虚假性在进行下一次调查时变得明显，因为下一次调查不会显示出前一次调查所指出的下降迹象，但在最新调查之前发生了新的下降。关于这一现象的突出的例子，见图1。

### (b) 年龄选择误差

556. 在像通常所做的一样只对50岁以下的妇女收集生育史时，被排除在外的是调查前一年49岁或49岁以上的妇女的生育数，调查前两年48岁或48岁以上的妇女的生育数，依次类推，对于调查前各连续年份都是如此。为了直观地了解这种渐次的不选择更年轻妇女的影响，可以想象一张普查前35年的按一岁分组的分年龄组出生率表格。该表的大致样式见表9。

557. 这张表中的单元格中的比率表示调查前35年的生育率水平和趋势。只有从左上到右下延伸的单元格对角线以上的比率可以根据调查时年龄在50岁以下的妇女的生育史计算得出。对角线以下的单元格中的比率表示调查时年龄在50岁以上的妇女的生育经历。

558. 选择50岁以下的妇女对调查之前几年的生育率水平影响不大，因为相对而言将近50岁的妇女的生育数很少。在计算调查之前一年总生育率所需的生育数中，只有年初49岁以上的妇女的生育数被省略。对调查之前三年计算总生育率时不会考虑调查之前三年47岁或47岁以上的妇女的生育数。

559. 然而，年龄选择确实限定使用生育史数据计算生育率趋势。不为调查时50岁以上妇女估计缺失的生育率数值就无法计算总生育率趋势。可以

为调查之前10年计算40岁以下妇女的生育率趋势，为调查之前20年计算30岁以下妇女的生育率趋势，依次类推。

表9. 根据生育史获得的生育数

母亲年龄	调查前……年													
	35	34	33	32	31	30	.	.	.	5	4	3	2	1
15	*	*	*	*	*	*	.	.	.	*	*	*	*	*
16		*	*	*	*	*	.	.	.	*	*	*	*	*
17			*	*	*	*	.	.	.	*	*	*	*	*
18				*	*	*	.	.	.	*	*	*	*	*
19					*	*	.	.	.	*	*	*	*	*
.						.	.	.	.	.	.	.	.	.
.						.	.	.	.	.	.	.	.	.
45										*	*	*	*	*
46											*	*	*	*
47												*	*	*
48													*	*
49														*

注：\* =可能填写的单元格  
 . =继续  
 空白 =没有登录

## 六、人口普查和调查的死亡率数据

### 导 言

560. 本章探讨了可以列入人口普查和抽样调查以得出死亡率数据的问题。在几种情况下，用一组相关的问题获得生育率和死亡率数据。关于平均生育数和存活子女数的问题便是一个重要的实例。因此，第五章和第六章应该放在一起阅读。

561. 人口普查和调查中的两个差别对于收集死亡率数据来说是至关重要的。第一个差别是死亡率水平和趋势与死亡率年龄模式。本章仅讨论了五种方法中的两种，即生育史和近期住户死亡情况，这两种方法提供了关于死亡率年龄分布的信息。其余三种方法只是根据关于年龄分布的假设提供了死亡率高低和趋势的估算数。根据关于死亡率年龄分布的数据来检验这些假设永远都是理想的做法。

562. 第二种差别是年幼者，包括婴儿、儿童和青少年的死亡与所有年龄的成年人的死亡。本文只讨论了五种方法之中的一种，即近期住户死亡情况，这种方法提供了所有年龄的死亡率信息。关于平均生育数和存活子女数的问题主要提供了15岁以下的人的情况，但在某些情况下，上限可高达30岁。生育史只提供15岁以下的人的信息，样本规模一般将有用的信息限制在5岁以下的人。关于父母和兄弟姐妹存活情况的问题只提供成人死亡率的信息。

563. 本章讨论的所有方法都可以用来补充死亡登记数据，特别是可以用来提供死亡登记数据中无法识别的人口亚群的死亡率的信息。但是，当没有提供可靠而完整的死亡登记数据时，至少采用本章所述的五种方法中的四种一般来说才是明智的，尽管在相同的数据收集工作中不一定要全部使用。关于平均生育数和存活子女数的问题可用于提供婴儿和儿童死亡率高低和趋势的估计数。生育史问题可用于提供婴儿和儿童死亡率的年龄分布的估计数。关于近期住户死亡情况的问题可用于估计成人死亡率年龄分布，关于父母和/或兄弟姐妹存活情况

的问题可用来估计成人死亡率高低和趋势。

564. 不完整的死亡登记数据将提供关于死亡率的年龄分布情况，这种信息可以与本章中讨论的人口普查和调查办法放在一起使用，但是，必须认识到，死亡登记是否完整可能是年龄选择问题，特别是，婴儿和幼儿的死亡报告可能不如老年人死亡报告那么充分。

#### A. 出生婴儿的存活情况

565. 在过去50年，全世界都使用关于平均生育数和存活子女数的问题来收集关于婴儿和儿童死亡率的数据。这些问题总体上非常成功，可以用在所有类型的人口普查和调查中。第六章A节应该与第五章D节“平均生育数”一起来阅读。

#### 1. 问题

566. 第五章D节第1小节中已经详细提出了关于平均生育数和存活子女数的问题。这些问题旨在查明被调查的每个妇女的以下情况：(a) 该妇女一生中总共生育的女孩数；(b) 该妇女一生中总共生育的男孩数；(c) 存活的女孩数；和(d) 存活的男孩数。

567. 一般都应该得到出生男女婴儿数及存活女孩数和男孩数。得到的数据既可用于单独估计男女两性的死亡率高低和趋势，又可用于内部评估数据质量。但是，如同在某些人口普查情况下，如果限制问题数目方面的压力过大，可将问题限制为平均生育总数和存活子女总数。

568. 从平均生育（男女）数中减去存活儿童（男女）数，便得出死亡儿童（男女）数。死亡儿童数除以相应的平均生育数，便得出死亡比例，从中可以推算出测定婴儿、儿童和青少年死亡率的生命表测度。

## 2. 制表

569. 除了用“存活儿童数”代替“出生婴儿数”一项之外，儿童存活的标准表与出生婴儿数据的标准表相同。但是，当数据用于估计婴儿和儿童死亡率时，使用单独表格会使计算出来的死亡儿童比例出现偏差。为了避免这种偏差，可以采用本节所述的专用的综合制表法。

570. 处理“不确定”情况的标准程序是在其所发生的任何表格栏目中加入“不确定”空格。但是，如果平均生育数和存活子女数的单列表格都采用这一程序，报告出生子女数但没有报告存活子女数的妇女将被视为其所有子女都已死亡。这将使计算出来的死亡儿童比例升高。报告存活子女数但没有报告出生子女数的妇女将被算为其死亡子女数为负数。这将使计算出来的死亡儿童比例下降。

571. 为了避免这种偏差，计算死亡儿童占出生婴儿总数的比例必须仅仅针对就这两个问题均做出答复的妇女。表10所示的列表形式提供了这方面的一种办法。

572. 根据前三列计算死亡儿童占所有儿童的比例。右面三列仅用于诊断目的，并提供接受提问的妇女的全面情况。被调查的妇女人数就是第一列与最后三列之和。欲知详情和实例，见Feeney的论著，1976年。

## 3. 估计

573. 出生婴儿和存活儿童数据可用于估算人口普查或调查之前二十年左右的死亡率水平和趋势。最初的估算程序假设死亡率不变，且只提供死亡率高低的估计数。后来的程序允许估算死亡率趋势，但是死亡率水平没有大的变化。《手册十：人口间接估计技术》（联合国，1983年，第3章）讨论了估算程序。

574. 存活儿童的信息仅仅用于估算成年女性的死亡率（Preston, 1980年）。此项技术不如基于出生婴儿和存活子女数据的那些方法有效，因而没有被广泛采用，但这种方法具有两个宝贵的特点。第一，因为这种方法不要求提供平均生育数，所以不受死亡儿童数漏报的影响。第二，如果存活儿童数有漏报现象，这种方法往往夸大死亡率高低。因为大多数估算死亡率的方法往往都低报死亡率水平，这种方法所产生的估计数与大多数其他方法所产生的估计数之间的一致之处，便是估计数准确的证据（如果a和b是c的估算数，且 $a \leq c$ ， $b \geq c$ ，那么， $a=b$ 则表示 $a=b=c$ ）。

## 4. 质量

575. 经验表明，在各种数据收集工作包括人口普查中，关于平均生育数和存活子女数的问题还是有可能获得高质量的答复的。如果出现错误，往往也是低估存活儿童数和出生婴儿数。但是，当

表10. 用于估计死亡率的出生婴儿和存活儿童表

母亲年龄	报告CEB和CS <sup>a</sup> 的妇女	平均生育数 <sup>b</sup>	存活子女数 <sup>c</sup>	报告CEB但没有报告CS的妇女	报告CS但没有报告CEB的妇女	没有报告CEB或CS的妇女
15-19	*	*	*	*	*	*
20-24		*	*	*	*	*
•						•
•						
80-84						
85岁以上						

注：\* = 可能填写的单元格

• = 继续

<sup>a</sup> CEB = 平均生育数；CS = 存活子女数。

<sup>b</sup> 不包括没有报告存活子女数的妇女的平均生育数。

<sup>c</sup> 不包括没有报告平均生育数的妇女的存活子女数。

数据用于估计死亡率时，起作用的还是存活儿童的比例。如果平均生育数和存活子女数都被低报，就会发生某些遗漏错误。如果平均生育数和存活子女数被低报到相同程度——例如，如果两者被低报10%，计算出来的存活和死亡比例都将是正确的。

576. 实际上，如果漏报成了一个问题，存活子女报告可能比平均生育数报告更加完整。计算出来的死亡儿童比例因此也可能过低而不是过高。如果以这种方法得出的死亡率估计数中存在着严重错误，则估计数一般过低而不是过高。

577. 但是，不完善的报告不是根据这些数据产生的死亡率估计数的准确性的惟一影响因素。从平均生育数和存活儿童数推算死亡率估计数时，要求就死亡率的年龄分布和生育率的年龄分布做出假设。这种假设也可能导致误差，这些误差也会导致估计数过低或过高。对生育率的年龄分布的假设所导致的误差一般较小，但对死亡率的年龄分布的假设所导致的误差可能很严重。

578. 如果出身婴儿数和存活儿童数据获自不同时间点进行的两次或两次以上的数据收集工作，这些数据的价值则大增。比较各套数据的重叠部分便可对估计数的质量进行有力的检验。其要点与对第三章A节第4小节中的生育率估计数所做的检验相同。

#### B. 根据生育史得出的死亡率数据

579. B节论述了将生育史作为婴儿和儿童死亡率的数据来源问题。本节应与第五章F节“生育史”一起阅读，该节讨论的是将生育史作为收集生育率数据的方法问题。

580. 将生育史作为死亡率数据来源的广泛经验始于世界生育率调查方案。世界生育率调查产生了婴儿和儿童死亡率数据，这种数据主要作为获得准确的生育率数据的工作的副产品。但是，研究由此产生的数据为死亡率研究做出了重要贡献，利用生育史来收集死亡率和生育率数据，现已成为一种惯例（Preston, 1985年），特别是，成为人口与健康调查以及类似调查中的一项重要内容。

581. 除了主要登记所有婴儿和儿童死亡事件的民事登记系统以外，生育史可能提供了最准确的婴儿和儿童死亡率的年龄分布数据。这种信息可用于检

查根据出生婴儿数和存活儿童数数据产生死亡率估计数时所使用的假设的死亡率年龄分布是否有效。

#### 1. 问题

582. 第五章F节第1小节讨论了生育史问题。这些问题旨在查明被调查的每个妇女所生的每个子女的以下情况：(a) 每个子女的性别和出生日期；(b) 该子女是否活着；如果否，(c) 其死亡的年龄和/或日期。

#### 2. 制表

583. 表11是根据生育史数据计算婴儿和儿童死亡率的主表。因为调查访问一般是在几个月内进行的，所以表格中的时间一栏填的是访问之前几个月而不是历月。

584. 如果得到15-49岁——年龄跨度为35岁——妇女的完整生育史，则底部（死亡）共有420行，15至49岁（含49岁）每个月一行，另有420列，总共为176 400个空格。从右上方到左下方的对角线下面的空格一定为零，因为这些空格中应该填的是访问后发生的死亡，因此还有88 410个可能不是零的空格，表11中用星号（\*）来表示。这一般高于生育史中所记载的生育总数，所以比已经死亡的出生数高得多。因此，本表中大多数可能不是零的空格代表零或死亡数为1。如果样本未加权或自行加权，则表中的空格将几乎全部由零或1组成。在加权（而非自行加权）的样本中，空格可能包含分数。

585. 表11一般是计算机内部生成和处理的，因此汇总统计资料适合打印和进一步处理。即使曾经打印出来过，这种表本身也很少见。根据计算机处理的详情，只有本表某些组别的空格的总和在计算机内部中生成。但是，无论计算机处理的详情如何，表11都是根据生育史数据计算婴儿和儿童死亡率统计资料的逻辑基础。

586. 用图2中所示的列克西斯图来表示表11中所示的各种出生和死亡数字是有用的。访问前头一个月出生的儿童由图2中的bc线来表示，其数目在表11“出生”栏最右边的空格中列出。0周月（和调查访问之前）的儿童死亡由三角形A来表示，其数目在表11“死亡”栏右上方的空格中列出。同样，调



表11. 按出生月份（调查访问前几个月）分列的儿童和按出生月份（调查访问前几个月）和以周月表示的死亡年龄分列的死亡儿童

		出 生						
		出生月份 (调查访问前……月)						
		•	•	•	4	3	2	1
		*	*	*	*	*	*	*
		死 亡						
死亡年龄 (周月)		出生月份 (调查访问前……月)						
0		•	•	•	*	*	*	*
1		•	•	•	*	*	*	
2		•	•	•	*	*		
3		•	•	•	*			
•		•	•	•				
•		•	•					
•		•	•					

注：\* =可能填写的单元格  
 • =继续  
 空白 =没有登录

查前第二个月出生的子女数量由ab线来表示，其中0周月的儿童死亡由平行四边形B来表示，1周月（和调查访问之前）的死亡由三角形C来表示。因此本表各栏对应于图中的对角线。

587. 该图显示，表11提供了计算访问日期前第二个月及此前每个月年龄、达到任何足月的人的出生概率所必要的信息。例如调查前第二个月出生的儿童在满一个月之前死亡的概率是平行四边形B所代表的死亡数除以ab线所代表的儿童数。同样，访问前第二个月内恰好满1个月的儿童在满两个月之前死亡的概率为平行四边形D所代表的死亡数除以de线所代表的子女数。

588. 该图还显示，访问之前那个月出生的儿童的概率无法直接算出，因为分子中所代表的某些死亡是在访问后发生的。这些死亡由图表2中的年龄轴最右端的各点来表示。

### 3. 估计

589. 在任何特定期间内达到实足年龄 $x$ 的人在实足年龄 $x+n$ 时的生命表死亡概率被界定为这些人的人数除以在达到实足年龄 $x+n$ 之前死亡的人数。图

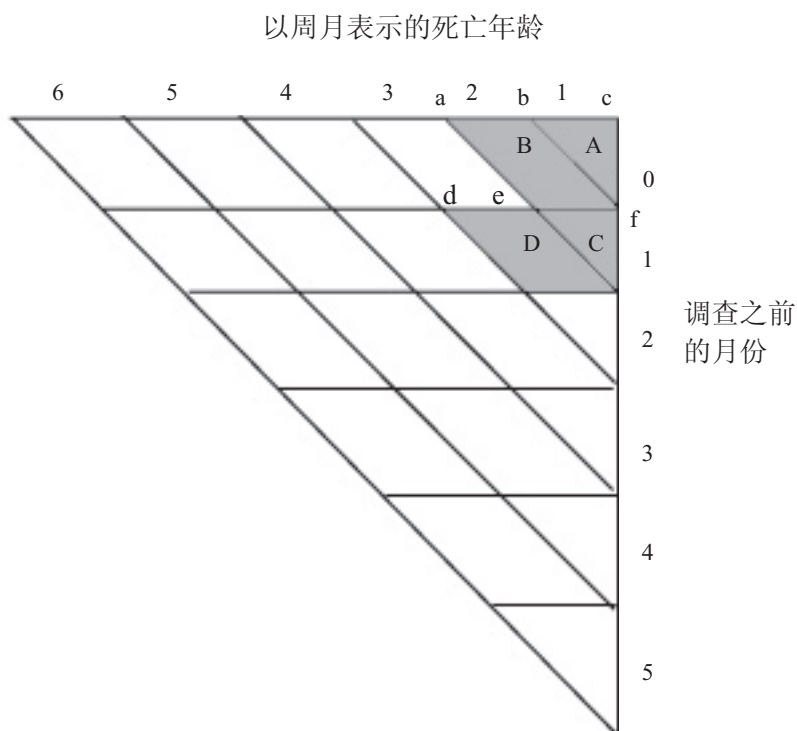
3中的列克西斯图说明了在特定调查前的一年中实足年龄满3个月的儿童在实足年龄6个月时的死亡概率的可行办法。其他年龄组和期间死亡概率也可以这样计算。

590. 对死亡登记数据采用的标准办法是：首先计算该年龄组和该期间年龄别死亡率，然后使用一个数学公式，将该死亡率换算成生命表概率。图3中A部分的矩形代表该比率的分子所计入的死亡数和分母中所计入的有这种风险的人年数。

591. 对生育史数据一般采用不同的方法，这种方法可以直接计算死亡概率。计算图3中B部分的bgid平行四边形所代表的期间和年龄组的死亡概率就是为了抵消平行四边形afje，即概率的分子是这一平行四边形所代表的死亡数，概率分母则是af线所代表的人数。当然，三角形hji所代表的死亡数无法获得，因为这些死亡是在访问之后发生的。因此有必要以某种方式粗略估计其数量。例如，可以通过三角形fgh所代表的死亡数或三角形ced所表示的数粗略估计其数量。

592. 计算分母时从对应的组群中的出生子女数开始并减去在达到特定年龄前死亡的子女数。回

图2. 根据生育史推算出的死亡率数据的列克西斯图示



到图2，例如，图2中的ef线所代表的访问前那个月满1个月的子女数可以这样计算：ab线所代表的对应组别的出生数减去平行四边形abfe所代表的0周月时死亡的活产数。

593. 计算死亡概率不必使用表11中的足月细节。图3的C部分说明了Sullivan、Rutstein和Bicego所述的程序（1994年，附件A）。他们建议估计bc线所代表的儿童数为ac线所代表的数量的一半、gh线所代表的儿童数为gi线所代表的数量的一半。bh线所代表的死亡概率的分母则为bc线、cg线和gh线所代表的儿童数之和。

594. 同样，作者建议估计平行四边形bcfe所代表的死亡数为平行四边形acfd所代表的死亡数的一半、平行四边形ghkj所代表的死亡数为平行四边形gilj所代表的死亡数的一半。平行四边形bhke所代表的概率的分子则为平行四边形bcfe、cgjf和ghkj所代表的死亡数之和。

595. 表12显示了根据表11的数据产生的婴儿和儿童死亡率的说明性汇总统计数字。计算该表第三部分中所示的死亡概率的方法是：把该表第四部

分中所示的死亡数除以该表第五部分中所示的有风险的人数。

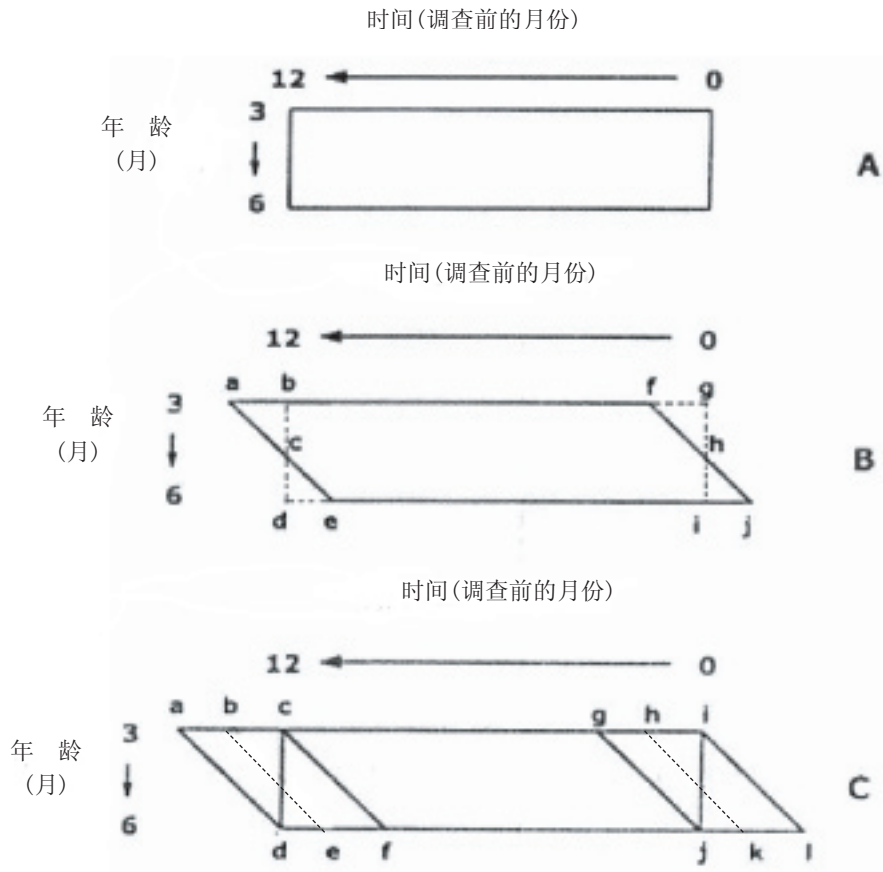
#### 4. 质量

596. 如果不能得到完整而准确的死亡登记数据，生育史数据一般可以提供获得高质量的婴儿和儿童死亡率数据的最佳机会。涉及的问题详细，访问员有可能深入调查，这些都有助于将报告的误差降低到最低限度。但是，为了使这些优势得到发挥，必须保证外地工作人员的培训和监督达到高水平。答问者能否准确回忆起事件的日期以及他们愿不愿意提到死去的子女，当然会对所收集数据的准确性有所限制。

##### (a) 误报死亡年龄

597. 死亡儿童的死亡年龄报告不准确是一个常见的问题。如果按周月报告死亡年龄，“年龄大量堆积”在12个月上的倾向可能导致某些婴儿的死亡被报告为是在周岁为1岁时发生的。如果对实足年龄2岁以前死亡的儿童包括婴儿期死亡的儿童的死亡年龄按死亡时的月龄制表，就可以发现这种堆积现象。

图3. 说明死亡概率计算方法的列克西斯图



(b) 年龄选择偏差

598. 只收集50岁以下妇女的生育史时，常见的情况是，根据生育史计算出来的追溯性统计数据中存在着年龄选择偏差。由于以前的统计数据局限于年轻妇女的经历，所以可能出现年龄选择偏差。

599. 年龄选择偏差如图4所示。调查之前35年的生育情况用一个正方形表示，该正方形由时间0至-35的时间轴向下延伸的垂直线与年龄15至50岁的年龄轴向左延伸的水平线构成。这里假设所有生育事件都发生这一年龄范围。

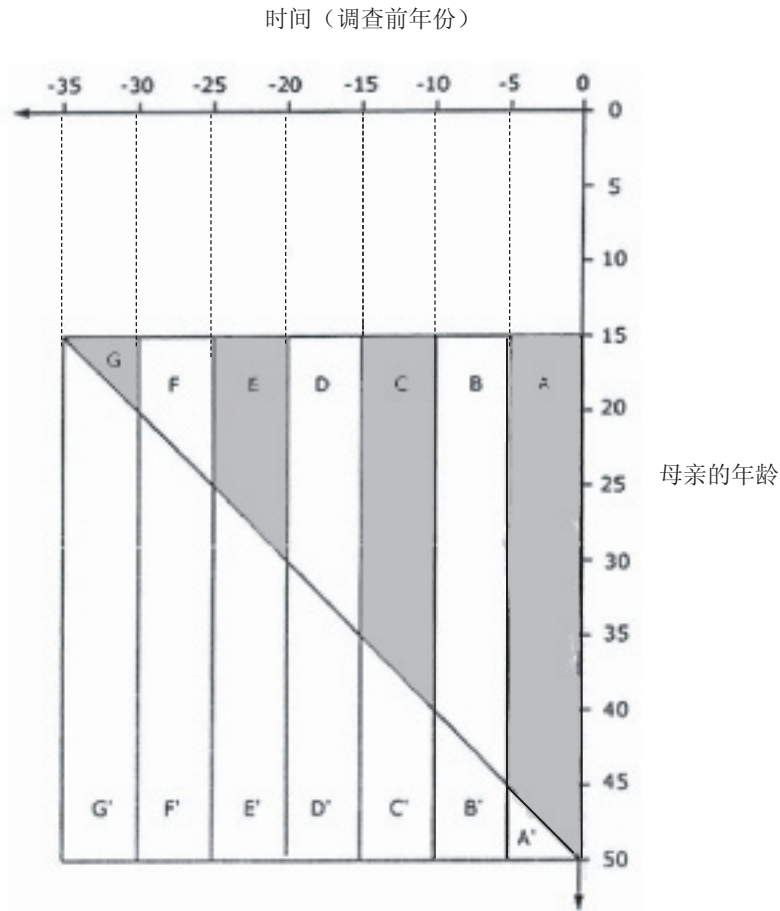
表 12. 缅甸根据生育史数据得出的婴儿和儿童死亡率统计资料

按以月表示的选定年龄间隔、婴儿死亡率、五岁以下死亡率和死亡概率分子和分母、选定期期分列的死亡概率，1991 年，缅甸				
男女两性				
调查前的年份 (年)	0-14 1976-1990	0-4 1986-1990	5-9 1981-1985	10-14 1976-1980
婴儿和五岁以下死亡率				
婴儿死亡率	100.1	102.9	96.1	101.6
5q0	138.2	146.9	126.5	140.6
以月表示的实足 死亡概率				
年龄				
0-1	55.9	53.5	54.6	60.3
1-3	16.9	18.9	15.8	15.9
3-6	14.2	16.3	13.5	12.6
6-12	16.5	17.9	15.3	16.1
12-24	13.1	14.6	10.7	14.2
24-36	9.1	10.4	6.4	10.9
36-48	12.3	14.9	11.2	9.7
48-60	8.4	10.2	5.7	9.4
以月表示的实足 死亡数				
年龄				
0-1	887.0	296.0	310.0	281.0
1-3	253.0	99.0	85.0	69.0
3-6	208.0	84.5	71.0	52.5
6-12	237.0	93.0	78.5	101.0
12-24	183.5	76.0	52.5	55.0
24-36	122.5	54.0	29.5	39.0
36-48	158.0	76.5	50.0	72.0
48-60	103.0	51.0	24.0	28.0
以月表示的实足 承受风险 (有风险的人数)				
年龄				
0-1	15 878	5 534	5 682	4 663
1-3	14 952	5 244	5 372	4 337
3-6	14 640	5 189	5 272	4 179
6-12	14 377	5 182	5 134	4 061
12-24	13 969	5 201	4 884	3 884
24-36	13 421	5 214	4 643	3 564
36-48	12 851	5 148	4 466	3 238
48-60	12 192	5 003	4 200	2 989

资料来源：根据《1991年人口变化与生育率调查》（移民和人口部，1995年，仰光）表7.1。

注：死亡概率的计算方式是死亡数除以有风险的对应人数。婴儿死亡率和五岁以下死亡率(5q0)是根据生命表法的死亡概率计算的，例如，调查前所有0-14岁儿童的婴儿死亡率为1减1-0.05587、1-0.01692、1-0.0142和1-0.01648的得数。

图4. 显示根据生育史数据计算的死亡率估计的年龄挑选偏差的列克西斯图



600. 连接图4中的正方形左上角和右下角的线将正方形分为两个三角形。右上的三角形代表调查时50岁以下妇女的生育情况。左下的三角形代表调查时50岁以上妇女的生育情况。50岁以下妇女的生育史显然不提供第二生育组的任何信息。

601. 对于调查前的五年期间来说，50岁以下妇女的生育史提供了A区所代表的生育情况（当然，没有活到调查之时的妇女在这一时期的生育除外），并且遗漏了A'区所代表的生育。因为在任何特定时期45岁以上妇女的全部生育数都只占较小一部分，所以遗漏这一部分的后果一般较小。

602. 但是，对于以前的各时期来说，遗漏更加严重，对于生育史能提供信息的最早时期，即截至调查前30年的五年期间来说，生育史只提供了G区所代表的生育数据，遗漏了G'区所代表的生育数据。G区所代表的生育只发生在20岁以下的妇女身上，一般说来发生在17.5岁以下的妇女身上。

603. 现在考虑一下这种年龄选择对计算调查前各个五年期间的婴儿死亡率的影响。在育龄期任何一端生育的婴儿死亡率往往高得多。20多岁妇女生育的婴儿死亡率往往最低，低于这个年龄和高于这个年龄的妇女生育的婴儿死亡率都比较高。因此，计算调查前五年期间婴儿死亡率时将三角形A'所代表的生育排除在外的话，将会使计算出来的比率降低，因为它排除了死亡率较高的一些生育。但是，这个偏差不会太大，因为靠近育龄期终端的生育事件比较少见。

604. 调查前各时期的偏差可能比较大，因为大部分生育都被忽略。但是，程度不能确定，因为尽管大部分生育被忽略（例如，B'区和C'区所代表的生育），但被忽略的生育与列入的生育的婴儿死亡率相差不大。

605. 对于最早时期来说，偏差方向发生逆转，程度可能严重。例如，调查前30至34年的婴

儿死亡率完全基于20岁以下妇女的生育，一般基于17.5岁妇女的生育。该比率一般比实际的婴儿死亡率，即根据该时期的所有生育计算的比率（G区加G'区）高得多。

606. 根据生育史数据计算婴儿死亡率的年龄选择偏差的重要性得到了广泛的研究或评估。只要计算的不是调查前15年以上的婴儿死亡率，偏差就可能不太大。

607. 由于离调查较远的时期的年龄选择偏差增大，所以可能出现虚假的时间趋势。只计算低于某一年龄的母亲生育的婴儿死亡率，就可以消除趋势偏差。例如，如果计算调查之前15年的比率，可能只会计算35岁以下妇女生育的婴儿死亡率。

608. 掌握50岁或50岁以上妇女的生育情况，就可以缩小年龄选择偏差。如果预期寿命相当高并且能够得到老年妇女的准确生育报告，就可能比传统上局限于育龄期的方法更加合适。

#### (c) 死亡率选择偏差

609. 由于在调查前各年生育的某些妇女在调查之前已经死亡并因此被排除在样本之外，所以出现了死亡率选择偏差。在一般情况下，由于死亡率相当低，生育史限制为50岁以下的妇女，并且计算调查前15年之内的婴儿和儿童死亡率，所以这种偏差有可能忽略不计。但是，如果上述任何条件都没有满足的话，死亡率选择偏差可能成为重要的因素。例如，如果为了缩小年龄选择偏差而提高收集其生育情况的妇女的年龄上限，死亡率选择偏差会更严重，因为年长妇女的死亡率会更高。例如，由于艾滋病病毒/艾滋病的肆意蔓延，育龄期的死亡率比平均死亡率高，这也可能导致严重的死亡率选择偏差。

#### (d) 抽样误差

610. 收集生育史的调查一般涉及5000至10000名妇女的样本。由于粗生育率为每千名妇女30名活产，所以调查10 000名妇女，将得到调查前各年每年约300名活产的资料。因此，在采访前的头一个五年里将有约1 500名活产。如果婴儿死亡率为每1 000名活产中100名婴儿死亡，预计婴儿死亡人数为150人。

611. 根据小部分变异系数约等于1除以死亡数的平方根这一近似规则，最终婴儿死亡率的变异系数约为8%。计算人口亚群时，抽样规模越小，或婴儿死亡率越低，产生的误差就越大。抽样设计因素经常使这一误差高达10%至15%。

612. 因此，将生育史作为死亡率数据的来源，涉及潜在的较低的非抽样误差与较高的抽样误差之间的抵消。关于近期住户死亡情况和父母、兄弟姐妹和平均生育数的问题可以列入大型调查之中，这种调查的抽样错差比列入全全生育史的大多数调查的误差少得多。这些问题也可以列入许多人口普查，其抽样误差为零。如果实地工作做得好，这些简单问题的非抽样误差可能与典型的生育史调查的抽样误差相近，或者小于后者。

#### C. 近期住户死亡情况

613. 关于近期住户死亡情况的问题不如关于平均生育数和存活子女数的问题使用得广泛。这部分上可能是因为早期的某些经验显示漏报死亡情况相当严重。但是，最近的经验表明，如果问题合适且实地工作细致，关于近期住户死亡情况的问题会产生好的结果。有鉴于此，这些问题可能比以往使用得更加广泛。

614. 关于近期住户死亡情况的问题具有若干重要的优点。如果没有完整而准确的死亡登记数据，这些问题则成为成人死亡率的年龄分布的惟一可能的信息来源。这些问题之所以重要，是因为提供成人死亡率数据的仅有的备选普查和调查问题对死亡率的年龄分布做出假设，只有提供了死亡登记数据或近期住户死亡数据，才能检验这些假设。

615. 近期住户死亡情况的问题也是可能提供整个年龄段死亡率数据的仅有的普查和调查问题。在某些情况下，存在婴儿和儿童死亡数严重漏报现象，但在另一些情况下，由于问题详细，调查细致，在婴儿和儿童死亡方面得到了可靠数据。

616. 关于住户死亡情况的问题尤其有利于获得关于死因的数据，特别是关于产妇死亡率的数据。任何普查或调查问题都不可能产生使用医疗证明的死亡登记系统所提供的高度详细而精确的死因资料。但是，如果补充问题适当而且实地工作细

致，关于近期住户死亡情况的问题可以产生有用的死因信息。

617. 即使在理想的条件下，也不能期望关于近期住户死亡情况的问题可以产生其完整性和精确性可与完善的死亡登记系统相媲美的数据。但是，对于许多国家来说，这些问题比现有的死亡登记系统能够产生更为完整的死亡报告。现在就有评估报告是否完整并且上调死亡报告数以更好地估算实际数字的方法。完整的报告永远都是最理想的，但并不要求为采用这些问题提出正当理由。近期住户死亡报告的完整率达到80%或以上的普查或调查通常将为测定死亡率做出重大贡献。

### 1. 问题

618. 关于近期住户死亡情况的问题与本章及上章讨论的其他问题不同，因为这些问题涉及到没有点查到或采访到的人。因此，他们的情况必须记录在问题调查表的专设栏里。这么做部分是出于这一原因，部分是由于必须得到关于每个死者的性别和年龄情况，所以关于近期住户死亡情况的问题可

以多种不同的方式措辞和确定格式。为了说明各种可能的情况，C节第1小节提出了四个实例。

619. 方框3列示了坦桑尼亚联合共和国1988年人口普查中使用的问题。这些问题旨在了解所有在普查日期前12个月内死亡的人的性别和死亡年龄，并保证填写总共多达3人的死亡情况。

620. 方框4给出了一种替代办法，即死者的性别和年龄以表格形式记录。明确列入方框3的问题和点查员指示在本方框中都是隐含的。但是，如果外地工作人员培训适当，这不会损害数据收集而且节省了表格的空间。

621. 方框3和4中的实例对人口普查前12个月内的死亡情况进行了调查。但是，也可以对更长时间的死亡情况进行调查，在这种情况下，一般可列入有关何时死亡这种补充问题。方框5列示了1990年中国人口普查中使用的问题。

该实例表明问题的详细措辞与数据收集工作的具体情形联系得多么密切。因为人口普查的基准时间是1990年年中，获得人口普查前18个月和这一时期半

方框 3. 近期住户死亡情况的问题，1988 年坦桑尼亚联合共和国人口普查						
F[节]. 住户死亡情况						
死亡发生情况	第一个人		第二个人		第三个人	
最近 12 个月家里是否有人死亡? 是=1 否=2  如果答案为否,请直接回答 第34 题	死者是男 是女?  男=1 女=2	死者死时多少岁?  写出周岁 数  如果不满 1 岁,请填写“00”	死者是男 是女?  男=1 女=2	死者死时多少岁?  写出周岁 数  如果不满 1 岁,请填写“00”	死者是男 是女?  男=1 女=2	死者死时多少岁?  写出周岁 数  如果不满 1 岁,请填写“00”
(27)	(28)	(29)	(30)	(31)	(32)	(33)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

方框 4. 近期住户死亡情况的问题，1988 年马拉维人口普查		
1987 年 9 月到现在家中发生的所有死亡情况：		
死者的性别？ M=1 F=2	死亡年龄？ (岁)	
C1	C2	
C3 死者总数 (如果无死亡，填写 00)		

注：M=男；F=女；C=问题代码。

年里的死亡信息提供了(a) 人口普查前一年里的死亡数，和(b) 1989历年的死亡数。

622. 最后这个实例表明，关于近期住户死亡情况和近期住户出生情况的问题如何可以结合起来以减少报告任一事件的遗漏机会的。这种方法主要优点在于它有可能改善婴儿和儿童死亡的报告情况。方框6列示了1991年缅甸人口变化与生育率调查中提出的问题。记录采访前12个月的出生和死亡情况的表格得到以下问题的补充：调查婴儿死亡情况的问题、关于婴儿死亡的死亡月份和日期的问题以及关于报告是否完整和所报告的出生和死亡情况是否一致的最后调查。

623. 这种方法对婴儿死亡情况的细致调查似乎非常有效，因为从这一数据中推算出的婴儿死亡率略高于从该调查中收集的生育史数据所推算出的比率（缅甸移民和人口部，1995年，第7章）。

## 2. 制表

624. 基本表将列示按死者性别和年龄分类报告的所有住户死亡情况。可以对整个人口编制按一岁分组的表格。如果年龄报告质量不佳，按以五岁分组的年龄组分列的表格对国以下各级地区和人口亚群都是足够的。但是，在每个表格中，0-4岁年龄组应该细分为零岁死亡（婴儿死亡）和1-4岁死亡。这些组别的死亡危险截然不同。如果不将它们区分开来，将造成大量信息损失。

625. 如果问题涉及更长一段时间的死亡情况或提供在基准期内何时死亡的信息，则应该将该信息列入表中。例如，就1990年中国人口普查中使用的问题来说，近期住户死亡情况表显示了方框5中的三个六个月期间的每个期间各个年龄的男女死亡人数。

方框 5. 近期住户死亡情况的问题，1990 年中国人口普查					
4. 住户死亡人数					
1989 年前 6 个月		1989 年后 6 个月		1990 年前 6 个月	
M _____ F _____ 总数 _____		M _____ F _____ 总数 _____		M _____ F _____ 总数 _____	
<input type="checkbox"/> <input type="checkbox"/>		<input type="checkbox"/> <input type="checkbox"/>		<input type="checkbox"/> <input type="checkbox"/>	



方框 6. 近期住户死亡情况的问题 1991 年缅甸人口变化和生育率调查

生育率表

死亡率表

最近 12 个月你家里是否有过活产，包括后来死亡的活产？

最近 12 个月你家里是否有人死亡？

请填写下表。

请填写下表。

序号	母亲		儿童性别	
	行号	年龄	男童	女童
(1)	(2)	(3)	(4)	
1			1	2
2			1	2
3			1	2
4			1	2
5			1	2
6			1	2

序号	姓名	性别		死亡年龄 (岁)
		男	女	
(1)	(2)	(3)		(4)
1		1	2	
2		1	2	
3		1	2	
4		1	2	
5		1	2	
6		1	2	

如果是双胞胎，请使用括号，每个孩子填写一行。

填入出生总数□

填入死亡总数□

1.(a) 最近12个月家里是否有婴儿死亡（一岁或一岁以下的儿童死亡）：

是……1（回答第 1(b)题）

否……2（回答第 2 题）

(b)

性别		死亡年龄（天或月）			
男	女			或	
1	2	天……	□□□	或	月…… □□□
1	2	天……	□□□	或	月…… □□□
1	2	天……	□□□	或	月…… □□□
1	2	天……	□□□	或	月…… □□□

2. 只是为了保证我填写的内容都正确：最近 12 个月家里是否有孩子出生后不久便死亡？

是……1

（访问者：检查并改正生育率表、死亡率表和第 1[a] 题中的错误）。

否……2

3. 估计

626. 关于近期死亡情况的问题提供了完整的民事登记系统可能收集的死亡信息的“替代”。如果报告完整（受到选择误差造成的限制），它们可能直接用作年龄别死亡率的分子。

627. 但是，一般应当假设报告不会完整，因此有必要估算和调整未报告的死亡情况。有关这

种做法，见Bennett和Horiuchi（1981和1984年）、Hill（1987年）和联合国的文献（1983年，第5章）。

4. 质量

628. 关于近期住户死亡情况的问题的早期经验通常导致遗漏程度高，但近期经验表明，在许多

情况下，提出这种问题可能产生好的结果。尽管期望完整的报告可能是不现实的，而且必须做出特别努力才能避免婴儿和儿童死亡事件漏报，但将来这种问题似乎比过去能够起到更大的作用。

#### (a) 答复误差

629. 常见的答复误差是死亡情况漏报。鉴于任何住户的死亡事件的性质和重要性，答复者不可能不知道死者的情况，也不可能在这这么短的时间里便将他们遗忘。就像近期出生情况一样，这一定另有原因。

630. 近期死亡事件漏报的原因可能与近期出生事件漏报相同。缺乏训练的点查员可能回避提问。答复者可能不愿意报告死亡情况，或者可能由于他们不能确定死亡事件的确切发生时间而不知道是否应该报告。前两个问题对于近期死亡来说比对近期出生可能更严重。点查员可能不喜欢询问死亡情况，因为答复者不喜欢回答有关死亡的问题。还有一个原因是，近期出生者一般都在家中——已经死亡的人则不然——这显然使实地工作人员和答复者想起最近发生的出生事件。

631. 人们还可以有把握地假设，近期死者的性别的报告准确，但其死亡年龄的报告可能不如住在家中的人的年龄那么准确。死亡年龄的准确报告与家中健在成员的年龄的准确报告同等重要。如果提出关于住户死亡情况的问题，同样应该努力地去获得尽可能准确的报告。

632. 婴儿和儿童死亡情况报告得可能不如老年人死亡那么完整。特别是婴儿死亡可能是在出生后不久发生的，孩子还没有机会成为家中的正式成员，而且幼儿死亡对家庭的影响可能不如大孩子和成人那么严重。但是，上述的缅甸调查表明，如果问题非常详细而且实地工作质量高，甚至有可能得到可靠的有关婴儿的调查结果。

633. 缅甸的实例证明了根据其他国家或其他数据收集工作的经验或根据基于该经验形成的一般意见而普遍化的危险。具体问题在具体环境中是如何发挥作用的，这种问题最好是通过检验而不是依靠“先见之明”来解决。

#### (b) 选择误差

634. 在答复有关住户死亡情况的问题时报告的死亡事件，必须是在人口普查或调查时已有的住户内发生的。不会报告人口普查或调查之前不复存在的家庭的死亡情况。这显然包括一人一户的家庭中发生的所有死亡情况。同样，成年家庭成员的死亡也可能成为家庭解体的原因，该死亡也可能被遗漏。这种选择误差可能在艾滋病病毒/艾滋病蔓延的情况下极其严重。

635. 这种遗漏就遗漏本身而言不如按年龄选择并因此歪曲了死亡报告的年龄分布的可能性那么重要，在某种程度上这种遗漏已经预料到而且能够调整。作为年龄选择根源的这种遗漏没有得到令人满意的研究，尽管关于住户结构的普查和调查数据为这种研究提供了基础。

#### D. 父母存活情况

636. 人口普查或调查时母亲或父亲仍然活着的答复者所占比例反映了成人死亡率水平。死亡率高，则存活比例低，死亡率低，则存活比例高。从这一信息中拟订出估算人口普查或调查之前各年成人死亡率水平和趋势的程序。从这种问题中推算出的数据通常被称为“孤儿法”数据，但是应该注意的是，这个问题会问到成人和儿童。

##### 1. 问题

637. 最简单的问题是：

你的（该人的）母亲是否还活着？

正是因为从这一问题推算出的数据，人们首先制定了估算女性死亡率水平的方法。后来，人们又根据有关男性的相应问题制订了估算男性死亡率的方法：

你的（该人的）父亲是否还活着？

当询问母亲或父母死亡的年月时，就可能详细提出这两个问题（Chakiel和Orellana，1985年）。如果能够得到这方面相当准确的信息，这些数据就简化了估算程序并提供了近期的死亡率估计数。

638. 一种要求不太高的替代做法是询问每个

已婚者以下问题：(a) 该人的母亲或父亲在其结婚时是否还活着；和(b) 该人的母亲或父亲现在是否还活着。也可以从中得到该人公婆（岳父母）的这些信息。

## 2. 制表

639. 关于母亲存活问题的基本表是按年龄、性别和在人口普查或调查时母亲是否存活分列的全体人口表。年龄按五岁分组对于所有表格来说都已足够，但应该避免分为大组。未定后限的年龄组可以被认为是60+（60岁以上）。

640. 建议列入性别栏，因为人们注意到，男性答复者的取向偏差（见下文）比女性答复者更明显。因此，如果可能按答复者性别编制估计数，则有助于评估取向偏差。如果存在答复偏差，男性和女性答复者的报告将提供近乎相同的存活母亲和父亲的比例。

## 3. 估计

641. 根据存活父母的比例估计成人死亡率的想法是简单的想法。例如人口普查或调查时20-24岁的人一般是在22.5年前出生的，其母亲在22.5年前是活着的。已知人口的平均生育年龄约为28岁，所以存活母亲的比例为28岁至28+22.5岁或50.5岁的生命表存活概率的估计数。

642. 基本估算程序根据每个按五岁分组的年龄组（5-9岁、10-14岁……）的男性、女性或男女两性所占比例估计生命表存活概率。没有采用0-4岁年龄组，因为该年龄组的死亡父母的比例一般非常低，而且会有不规则的波动（Brass, 1975年：67）。对较大年龄组估计成人的年龄间隔期更长的存活情况。从10-14岁和15-19岁年龄组的数据中得出了25至40岁女性存活的估计数，从45-49岁和50-54岁年龄组的数据中则得出了25至75岁存活情况的估计数。

643. 由于年长者是在很久以前出生的，所以其父母可能死亡的时间比年轻人的父母时间长。这表明，不同年龄组答复者的存活父母比例的数据包含（如果只是隐含的）人口普查或调查前几十年的成人死亡率趋势的信息。1980年代，发展了根据父母存活数据估计成人死亡率趋势的技术（Brass,

1985年）。利用这种办法，可以估算出人口普查或调查前大约15年的成人死亡率趋势（联合国，1983年，第四章）。

## 4. 质量

644. 根据父母存活数据推算出的成人死亡率估计数的质量取决于数据质量、估计程序要求的补充参数的准确性以及估计过程所用的模型生命表家庭的有效性。一般说来，评估质量的最有效方法是比较父母存活估计数与其他来源的估计数。

### (a) 误报父母存活状况

645. 为了准确回答关于母亲或父亲的存活状况的问题，答复者必须(a) 知道母亲或父亲的特征，(b) 知道他们是活着还是死了，以及(c) 准确报告这些情况。

646. 报告自己的母亲或父亲的存活状况的答复者一般都知道他们的特征和他们的存活状况。只有在特殊情况下一些人才不知道父母的特征。父母一方或双方必定是在这个人很小的时候去世的，将其养大的人必定没有让他们知道其生身父母的特征。这种人可能报告他们视为母亲和父亲的人的存活状况。这种报告可能是正确的，尽管所报告的对象不对。但是，如果由于生身父母一方或双方死亡，子女被收养或寄养，则往往会将死者误报为存活，从而夸大父母存活情况。

647. 答复者在报告家里其他人的存活状况时可能不太能够而且也不太愿意准确地报告。误报机会取决于家庭规模和复杂性，取决于答复者与其所报告对象的关系。如果住户基本上是由核心家庭组成的，答复者一般知道家中每个人的父母的存活状况。家庭越大、越复杂，误报的可能性也就越大，因为答复者不知道提出的问题的正确答案。在这种情况下，误报倾向不太明朗，但答复者如果不知道该人已死，则可能假定他还活着。这样的话，将会出现夸大父母存活情况的趋势。

648. 如果答复者是所报告的对象收养或寄养父母，他们因为认为自己才是“真正的”父母，因而往往会报告其本人的存活状况而不是生身父母的状况。如果生身父母还活着，这种报告就是

正确的，即使所报告的对象不对。但是，如果生身父母已死亡，将他们报告为活着就是误报。这种误报往往夸大了父母存活情况，但是通过培训外地工作人员，让他们认识到这种倾向并向答复者强调这个问题涉及生身父母，就可将误报降低到最低程度。

649. 答复者报告的是自己的父母还是别人的父母，可能是由于母亲或父亲住在别处，其死讯还没传达到答复者。这会导致将母亲或父亲报告为存活，但事实上他们当时已经死亡。这种误报往往会夸大父母的存活情况。如果由于答复者报告收养或寄养父母而不是生身父母的存活状况而发生误报，则误报的程度可能随年龄而异。报告其父母存活状况的人的年龄越大，其收养或寄养父母已经死亡的可能性就越大。如果该人的生身父母已经死亡——这可能成为收养或寄养的理由，年长者误报情况就比年轻人少。这种误报造成的误差被称为取向偏差。

#### (b) 误报年龄

650. 误报年龄，特别是年龄夸大，将会导致根据父母存活数据推算的成人死亡率估计数的误差。年轻答复者的父母一般也比较年轻，因此其父母存活的可能性更大。夸大年龄将导致存活父母比例过高，因此死亡率估计数过低。有关年龄误报的影响的进一步评论，见Blacker和Mukiza-Gapere的论著（1988年；3.2.27）及Dechter和Preston的论著（1991年）。

#### (c) 选择误差

651. 无存活子女的男女的死亡情况必然没有包括在第637段的问题答案中。这种遗漏可能使估计数出现偏差。例如，如果无子女的男女的死亡率较高，最后的估计数往往过低。存活子女较多的男女（如果报告准确而完整的话）将被所有子女报告。因此，子女较多的人的死亡情况比子女较少的人报告得更充分。如果大家家庭的父母的死亡率较高，最后的估计数也往往过高。

#### E. 兄弟姐妹的存活状况

652. 根据兄弟姐妹的存活状况推算出成人死

亡率估计数的做法与根据父母存活情况或生育子女数的估计数所做的推算大致相同。这些问题的使用频率甚至低于关于父母存活状况的问题，所以这些问题的使用方面的经验比较少。因此，除非实地试验确定这些问题适合使用，否则不可能将它们列入人口普查之中。但是，在许多成人死亡率数据较少或根本没有的国家里，这种问题可能被广泛用在大规模的户口调查和生育率调查之中。

#### 1. 问题

653. 兄弟姐妹存活状况的基本问题如下：

##### 出生的兄弟数

你（该人）总共有几个兄弟，包括已经死亡的兄弟？

##### 存活兄弟数

这些兄弟中现在还活着的有几个？

##### 出生的姐妹数

你（该人）总共有几个姐妹，包括已经死亡的姐妹？

##### 存活姐妹数

这些姐妹中现在还活着的有几个？

654. 这些问题可能要问到所有的人，但是人们一般认为儿童信息不重要（关于平均生育数和存活子女数的问题有可能提供更可靠的儿童死亡率信息）。鉴于这种问题对成人死亡率的意义，可能对最低年龄，如10岁或15岁以上所有的人提出这些问题。

655. 这些问题的缺点在于答复者可能不知道婴儿时期夭折的兄弟姐妹。例如，答复者可能不知道其母亲所生的子女中有一个或几个先于他或她出生的孩子在婴儿期内夭折。

656. 有一套替代的问题只涉及在婴儿期或以后存活的兄弟姐妹。

##### 存活到15岁的兄弟

你（该人）有几个兄弟活到15岁？

##### 现在存活的兄弟

这些兄弟中有几个现在还活着？

##### 存活到15岁的姐妹

你（该人）有几个姐妹活到15岁？

##### 现在存活的姐妹

这些姐妹中有几个现在还活着？

对15岁和15岁以上的所有人提问这些问题。

657. 但是，第656段中的问题在解决了一个难题之后又提出了另一个难题。如果答复者认为难以准确报告年龄时，他们就可能难以知道某人是否活到15岁。尽管如此，成人兄弟姐妹存活问题总的说来普遍受到欢迎，因为童年末期和成年早期的死亡率概率最低。

658. 一种更详细的办法是问及旨在获得“兄弟姐妹情况”的问题，就是要得到一份列出每个兄弟姐妹的性别、出生日期、目前是否存活以及如果否其死亡日期（或年龄）的所有兄弟姐妹的清单。这种问题已经列入某些人口与健康调查，作为测定母亲死亡率的工作的一部分。

## 2. 制表

659. 第653段中的兄弟姐妹存活状况基本问题的基本表说明了按五岁分组的年龄组分列的15岁以上的所有人的情况。

问及成年兄弟姐妹的问题时，用“15岁时存活”代替“出生”，用“现在存活”代替“存活”。但是，兄弟姐妹存活状况基本问题的表格形式（见图示）和成年兄弟姐妹问题的表格形式相同。

## 3. 估计

660. 这种表用于计算每个年龄组存活的兄弟姐妹的比例。比例然后被换算成生命表男女“存

活概率”的估计数。进一步计算确定某一时点的存活概率，从较大年龄组推算出更早时点的概率。这样就推算出成人死亡率水平和趋势。有关估计方法，见Hill和Trussell（1977年）及Timæus等（1996年）。

## 4. 质量

### (a) 答问误差

661. 上文已经提到答问者不知道婴儿期死亡的兄弟姐妹的存在问题。由于缺乏了解有时导致遗漏了死亡的兄弟姐妹，所以从这些问题中推算出来的存活兄弟姐妹的比例将会过高，从而导致成人存活估计数也会过高。

662. 成人答问者一般都知道当他们本人在15岁左右时他们有几个兄弟姐妹，尽管他们可能不确定死亡的兄弟姐妹的确切死亡年龄。不愿意承认已经死亡的兄弟姐妹的存在将导致隐含的死亡数字过低。在这方面，父母存活状况的问题具有一个优点：无法隐瞒其存活受到怀疑的人的存在。

663. 如同对待平均生育数和存活子女数的问题一样，在培训和监督外地工作人员以尽可能确保最准确的报告方面应当共同作出努力。

### (b) 选择误差

664. 得不到独生子女的死亡情况，因为死亡的独生子女没有兄弟姐妹报告事实情况。如果独生子女的死亡危险与非独生子女不同，则将产生偏

年龄	人	在此居住	兄弟住 在别处	死亡	在此居住	姐妹住 在别处	死亡
15-19	*	*	*	*	*	*	*
20-24	*	*	*	*	*	*	*
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
80-84	*	*	*	*	*	*	*
85岁以上	*	*	*	*	*	*	*

注：\* =可能填写的单元格  
• =继续

差。但是，如果生育率适中或较高，只有一个子女的妇女的比例就小，成为独生子女的人所占比例更小，所以即使差异很大也产生不了巨大的影响。

665. 另一方面，兄弟姐妹多的儿童会被所有

存活兄弟姐妹报告为存活；因此，其死亡率情况会对结果将产生更大的影响。关于兄弟姐妹数量的信息可能有助于遏制这种影响。但是，人们一般都认为按存活兄弟姐妹数目分列的死亡率差异不可能大到造成严重偏差的地步。

## 七、 结论

666. 收集生育率和死亡率数据的方法和使用这些方法所引起的问题随时间而不同。人口研究导致了新的方法的产生。世界各国对新方法和现有方法的应用增进了对不同方法在不同环境下发挥什么作用的了解。技术发展为应用这些方法创造了新的机会。各国不断变化的经济、政治和社会条件改变了数据收集发生的环境和用户的需要、利益和复杂性。本手册试图选出一些基本原则，并对写书时的知识状况加以总结，但负责收集数据的组织和个人必须力求跟上该领域新的发展。这就是本章的重点。

### A. 人口研究

667. 收集生育率和死亡率数据的方法长期以来一直很稳定。本手册讨论的方法大多数都是十多年以前制定的，有很多已经使用了数十年。不过，新方法将不断出现，认识到这些发展是很重要的。各种会议、不同国家全国性数据收集组织之间的非正式联系，以及人口研究文献都是有关新的发展动态的信息来源。

668. 全新的方法可能较少，但有关现有方法在不同环境中的使用经验却在不断积累。实际应用现有方法所获得的知识在实践中可能比发展新方法更为重要。许多这样的知识都是在数据收集组织内获得的，并由于这些组织之间的联系而得到了最有效的共享。国家、区域和国际组织在这些知识的传播中可以发挥重要的作用，而人口研究文献往往对这些知识不够重视。

### B. 技术发展

669. 现代世界中的技术发展、特别是信息技术发展的速度之快是众所周知的。这对信息是其惟一或主要产品的数据收集组织产生了深刻的影响。过于强调信息技术在数据收集中的作用是错误的，因为有许多基本重要性它并没有提到。有效管理数据收集工作可能就是最重要的实例。同时，数据收

集组织工作的几乎每一个方面都无不受到信息技术发展的影响。因此关注这些发展动态是很重要的。

670. 近几十年来最重要的发展也许是互联网的出现，它是全世界各种组织和个人共享知识和信息的一种基本资源。互联网可能会成为不仅跟上信息技术的发展，而且跟上许多其他发展的一种基本工具。

### C. 用户和用户需要

671. 现代社会的日益复杂和日益相互依赖只能增加对信息的需求以指导各种行动，包括政府、企业、非政府组织、家庭和个人的行动。经济的全球化，以及社会和政治活动在一定程度上的全球化，将使各种决策者和计划人员更多地考虑全球或区域形势，而不是世界大约200个不同国家的形势。这将给统计定义和格式的适当统一以及全球和区域数据的汇编和表述带来越来越大的压力。

672. 数据和统计资料的用户当然是关于其需要和所关注问题的主要信息来源。情况始终如此，但是社会的快速变革和用户数量与多样性的增加，使得数据收集组织系统发展并保持与用户的双向联系以监测其不断变化的需要和所关注的问题变得更加困难，也更加重要。

673. 数据收集者往往从数据收集方法去考虑，并根据这些方法建立组织。用户则往往从其需要的信息或认为其需要的信息方面来考虑，而不太注意这些信息产生的方式。这两种观点都是正确而必然的，但二者之间存在着固有的紧张关系。由于用户需要是产生数据的最终动力，所以数据收集组织有责任对用户作出反应。当然，在有些情况下，恰当的反应也许是努力让用户了解数据收集的实情。

### D. 数据收集环境

674. 各种社会数据的收集以其发生的社会环境为条件。数据收集组织必须对影响数据收集社

会进程的社会环境变化保持敏感。社会环境丰富多样，因而不可能一一列举这些条件，但是一些影响的实例将会暗示必须予以关注的各种因素。

675. 不断变化的卫生条件可能需要对数据收集活动采取另外的方法并对这些活动提出新的要求。例如，艾滋病病毒/艾滋病流行病就改变了高发生率国家的人口模式，使人口估计的许多标准方法都不能用。

676. 国际人口流动，包括劳动力迁移和难民流动，在世界许多地方变得越来越重要。各国之间日益加强的这种相互依存对传统上不得不首先处理国家关注事项的数据收集组织提出了新的挑战。

677. 公民对隐私和政府和企业采取的数据收集程序之实际或感觉到的侵犯性的关注已经对一些国家的数据收集产生了重要的影响。这些关注经常影响将特定问题纳入人口普查和调查的决定。在若干极端情况下，它们使传统人口普查的使用变得无效。

678. 从技术和统计的角度来看，来自不同来源的关于个人的不同信息之间的联系丰富了所有来源，而国家授权的个人身份证号系统是有效实现这一联系的重要工具。不过，从更广泛的社会和政治方面来看，这些发展可能会对公民的隐私权构成威胁。这一问题对数据收集组织而言特别重要，因为没有公众的合作它们就无法工作。

## E. 国际合作

679. 与统计信息的收集和使用有关的事项方面的国际合作在当今很重要，在将来会更加重要。与过去和现在相比，在将来，使用某个国家收集的统计信息的用户会更多地设在其他国家和国际组织。这将增加对统计信息内容、其表达方式和它所描绘的“元数据”的国际标准化的要求。此外，从纸张信息到电脑信息的转变产生了需要加以解决的新标准化问题。由于这些原因，各国数据收集组织之间的通信与合作将变得更加重要。

680. 国际合作对数据收集组织的内部运营而言也是重要的。新信息技术创造的机会和风险及其到来的速度将使不同国家组织之间的经验共享变得前所未有的重要。目前有各种国际和区域会议和组织在促进这种通信和合作。它们在未来可能会变得更加重要。电信和电脑网络的发展将创造“远程交互作用”的新机制，以此作为面对面会议的补充。

681. 不同国家数据收集组织之间的双边和多边联系也非常重要。很少有人像他国从事同一工作的人那样了解数据收集的压力、风险和问题。由于这一原因，而且从其使数据收集组织的工作人员了解与数据收集有关的新发展的重要性方面来说，将数据收集者聚集在一起的国际会议非常重要。国际会议提供的面对面交流将在不远的将来发挥重要的作用，但可能会得到通过互联网实现的远程联系的日益补充和加强。



## 附 件

### 列克西斯图

682. 列克西斯图是观察各组人和事件及其关系的一种方法。它们提供了一种就人口制表、计算和统计进行交流 and 了解的专门而高效的语言。第六章B节运用列克西斯图来说明如何根据生育史数据计算死亡概率。本附件对列克西斯图进行了系统的说明，并说明了它们与第五和第六章讨论的其他方法的共同使用。

#### A. 列克西斯图表示法

683. 在一个以时间和年龄为轴的坐标面中，每一个人口事件可用时间和年龄坐标的交叉点来表示，时间为事件发生的时间，年龄为与事件有关的个人的年龄。每个人都可用直线来表示，称为生命线，该生命线将表示个人出生和死亡的点连接起来。

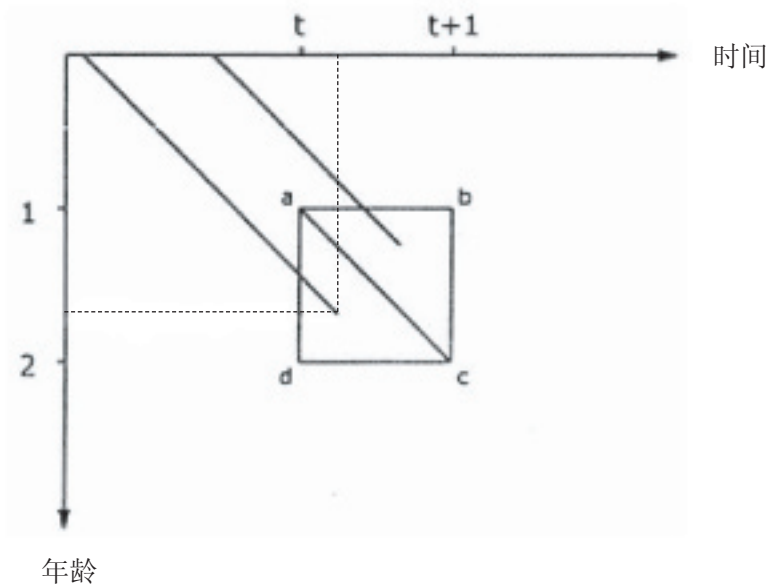
684. 图A1是两个人的生命线，二者均死于 $t$

年，周岁为1岁。这两条直线右下方的点代表直线所代表的个人的死亡。

685. 生命线提供了表示几组人的基础。通常的规则是，以坐标平面中的任何一条线代表其生命线与此线相交的一组人。例如，图A1中的 $ad$ 线代表在 $t$ 这一时点上实足年龄在1岁和2岁之间的人，直线 $bc$ 代表 $t+1$ 时点上实足年龄在1岁和2岁之间的人， $ab$ 则代表以 $t$ 开始的年份中实足年龄达到1岁的人。

686. 用点来表示事件提供了表示几组事件的基础。通常的规则是，可将任何一个二维图视为代表其表示点位于该图的一组事件。图A1中的正方形 $abcd$ 代表 $t$ 年（ $t$ 和 $t+1$ 之间的时间坐标）死亡的周岁为1岁（实足年龄1岁和实足年龄2岁之间的年龄坐标）的个人。

图A.1. 列克西斯图表示法



687. 由于代表个人的生命线向上延伸至代表此人死亡的点的左方, 图A1中的三角形acd代表t年初周岁为1岁的人的死亡, 三角形abc则代表该年达到实足年龄1岁的人的死亡。根据死亡人是(a) 属于期间初期的年龄组还是(b) 达到期间年龄组的下限, 对一年中达到某一特定周岁的人的死亡进行的这种划分对许多人口计算而言都很重要。图A1中两个三角形代表的死亡数是正方形代表的死亡数的一部分, 这被称为分离系数。

688. 图A1表示了几组死亡, 这些表示法原则也适用于各种人口事件, 如视为发生在母亲身上的事件的生育, 还有结婚和离婚等。该图没有显示代表的是哪种事件。必须根据实际情况补充这一信息。

689. 可将列克西斯图中各轴的不同方向用于不同的目的。例如, 可将时间轴画在下方, 年龄轴画在右方, 或者将年龄轴画在上方而不是下方。图A1中所示的方向如果不是使用最广泛的话, 也是使用最普遍的一种图, 因为它与出生和死亡的制表方式相应, 将代表较大年龄发生的事件的各行放在了代表较小年龄发生的事件的各行下方。

690. 年龄和时间可由代表时间流逝的不同变量代替。例如, 年龄轴可由代表结婚、第一次生育或*i*次生育以来时间变迁的轴来代替。时间轴通常代表日历时间, 但表示调查数据的图也可表示调查前的月或年等时间。

691. 列克西斯图被用于说明定义和关系。它们并不表示特定人口的定量方面。从来没有画过代表出生、死亡或某一特定人口内其他事件的所有的点, 除非是以一个非常小的人口进行举例。代表个人的生命线也不例外。

### B. 极端法

692. 有两种方法可以协助绘制和解释列克西斯图。极端法包括确定极端情况, 绘制代表这些情况的线或点, 并将这些线或点连接起来, 得到所需要的图示。

693. 例如, 在图A1中, 用直线ab表示以时间*t*开始的年份中达到实足年龄1岁的人。该例中的极端情况是与经过*a*点的生命线相应, 在*t*这一时点上

达到实足年龄1岁, 以及与经过*b*点的生命线相应, 在*t+1*这一时点上达到实足年龄1岁。列克西斯图的图示是将这两点连接起来的直线ab。

694. 图A3中的四边形abcd代表以时间*t*开始的年份中在*t*点上达到*x*岁的人的死亡。这里的极端情况是在时间*t*达到实足年龄*x*岁的人的死亡(*a*点), 在时间*t+1*达到实足年龄*x*岁的人的死亡(*b*点), 在时间*t+1*达到实足年龄*x+1*岁的人的死亡(*c*点), 以及在时间*t*达到实足年龄*x+1*岁的人的死亡(*d*点)。列克西斯图的图示是以这些点为顶点的四方形。

### C. 交点法

695. 交点法只适用于各组事件。它包括确定有关年龄组、期间和/或同期出生群, 画出年龄组、期间和/或同期出生群的图示, 并将这些图示的交点连接起来, 得到所需要的图示。

696. 同样以图A1中的正方形abcd为例, 该正方形代表以时间*t*开始的年份中死亡时周岁为1岁的人的死亡。它规定了某一期间, 即*t*年, 和某一年龄组, 即周岁1岁。

697. 这一期间发生的死亡以从时间*t*和*t+1*往下延伸的垂直线所形成的矩形区域表示。周岁为1岁的人的死亡以时间轴上经过这些年龄的水平线所形成的水平矩形区域表示。这两个区域的交叉部分, 即四方形abcd就是所需要的图示。

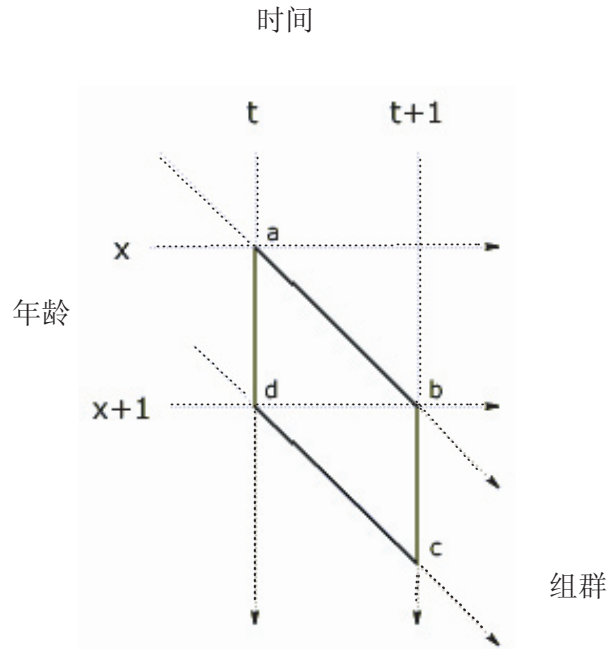
698. 以*t*年发生在*t*时周岁为*x*岁的人身上的后来事件为例(图A2)。该图确定了某一期间即*t*年和某一组群, 即在*t*这一时点上周岁为*x*岁的人。

699. 这一期间发生的事件以从时间*t*和*t+1*往下延伸的直线所形成的矩形区域表示。发生在这一组人身上的事件以平行对角线ab和dc形成的对角区域表示。两个区域的交叉部分即平行四边形abcd就是所需要的图示。

700. 最后以实足年龄*x*岁与*x+1*岁之间发生在从时间*t*开始的年份中实足年龄为*x*岁的人身上的事件为例(图A3)。该图确定了*x*岁和*x+1*岁之间的年龄间隔和从时间*t*开始的年份中达到实足年龄*x*岁的一组人。

701. 在年龄间隔期间发生的事件以实足年龄

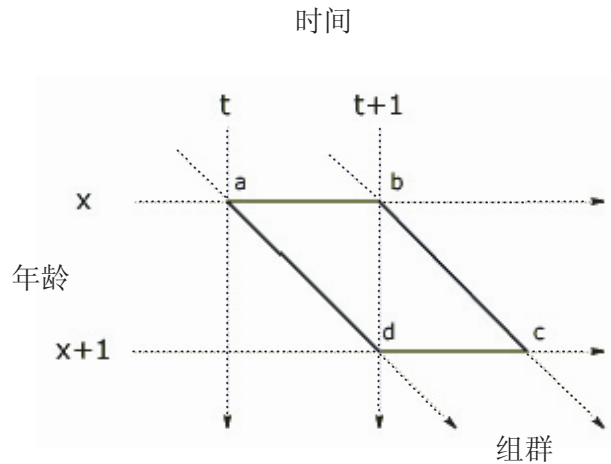
图A. 2. 期间-同组人事件



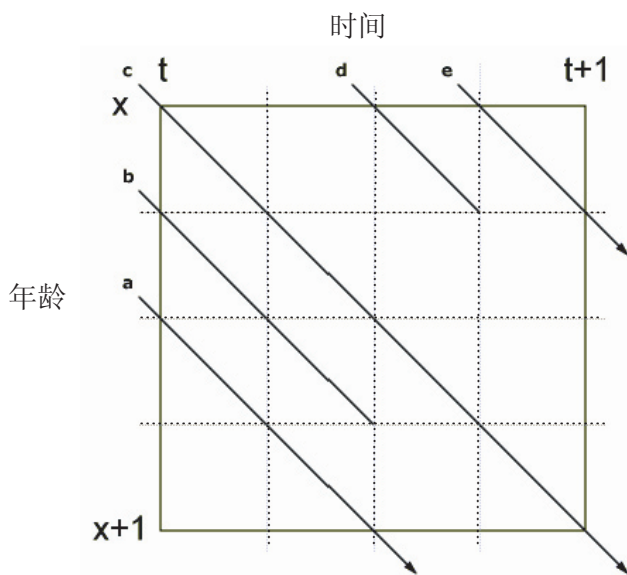
$x$ 岁和 $x+1$ 岁所在的直线形成的矩形区域表示。发生在这一组人身上的事件以延伸线段 $ad$ 和 $bc$ 的平行对角线形成的对角区域表示。两个区域的交叉部分即平行四边形 $abcd$ 就是所需要的图示。

702. 这两种一般方法虽然对绘制和理解列克西斯图有益，但只有熟悉它们在许多不同的实际情况下的运用才能最灵活地使用这些图示。以下各节列举了摘自手册主要部分的一些例子。

图A. 3. 年龄-同组人事件



图A. 4. 年龄别死亡率



D. 年龄别死亡率

703. 年龄别死亡率通常定义为某一特定期间某一特定年龄组的死亡人数除以这一期间这一年龄组人口中度过的人年数。分子中的死亡人数很容易解释。分母中的人年比较复杂，但通过图A4的说明就会变得很清楚。

704. 图A4给出了一个只涉及五个人的简单例子，分别以生命线a, b, c, d, e表示。a在时间t的实足年龄是 $x+1/2$ 岁，在 $x+1$ 岁后仍然活着。这个人在图中所示的期间和年龄组中活了 $1/2$ 年，给年龄别死亡率的分子增加了 $1/2$ 人年。b在时间t的实足年龄是 $x+1/4$ 岁，在 $x+3/4$ 岁时死亡。这个人也使图中所示的

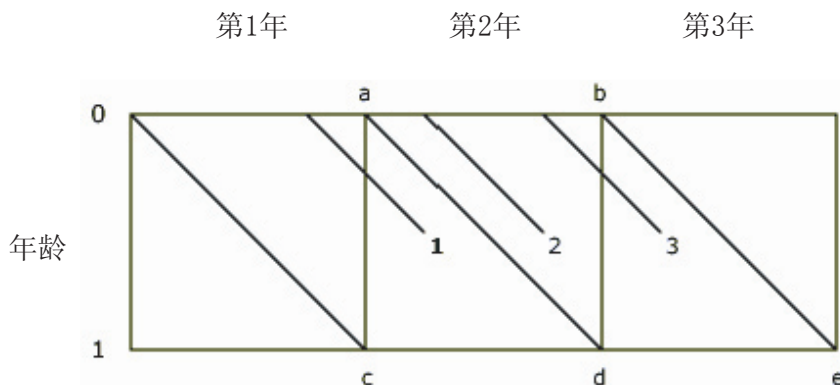
期间和年龄组的年龄别死亡率分母增加了 $1/2$ 人年。

705. c在时间t的实足年龄是x岁，在 $x+1$ 岁后仍然活着。他或她使年龄别死亡率的分子增加了1个人年。注意这是任何人可以增加的最大值。d和e给死亡率分母各增加了 $1/4$ 人年。

706. 年龄别死亡率的分子是这一期间任何时间的年龄组中仍然活着的每个人增加的人年之和。在这个例子中，五个人共增加了2.5个活过的人年。这一期间在该年龄组中有两人死亡，因此死亡率是 $2/2.5=$ 在度过的人年里每千人中有800人死亡。

707. 当人数和死亡数非常大时，用作年龄别

图A. 5. 婴儿死亡率



死亡率分母的活过的人年数可以大约计算为特定年龄组在特定期限初期和末期的平均人数乘以期间长度，这种计算不需要关于个人实足年龄和死亡年龄的数据。然而，当人数和死亡数较小时（调查数据经常如此），可以使用更精确的计算。

#### E. 婴儿死亡率

708. 婴儿死亡率通常定义为一个历年中婴儿死亡（实足年龄1岁以下的儿童死亡）人数除以该年的出生人数。这种计算方法并没有尊重风险承受原则，因为某个历年中发生的婴儿死亡事件中有些是上一年出生的婴儿的死亡事件，而某个历年中出生的有些儿童的夭折却发生在下一年。这一点在图A5中很明显。

709. 对婴儿死亡率的传统算法是用正方形  $abcd$  代表的死亡数除以直线  $ab$  代表的儿童人数。不过，在图中所示的三条生命线中，只有一条反映了存在风险的事件之间的正确关系。生命线1代表的儿童在第2年死于婴儿时期，但是出生于第1年。该儿童的死亡被列入死亡率的分子，但出生没有列入分母。生命线3代表的儿童出生于第2年，但在第3年死于婴儿时期。该儿童被列入死亡率的分子而不是分母。

710. 不过，死亡和承受风险之间的不匹配程度并不像本图所示的那样，因为几乎所有的婴儿死亡都发生在出生的头一个月，也就是说，都集中在代表实足年龄0的水平线附近。第2年出生的死于婴儿期的儿童大多在头一个月死亡，因此大多数将在第2年死亡。

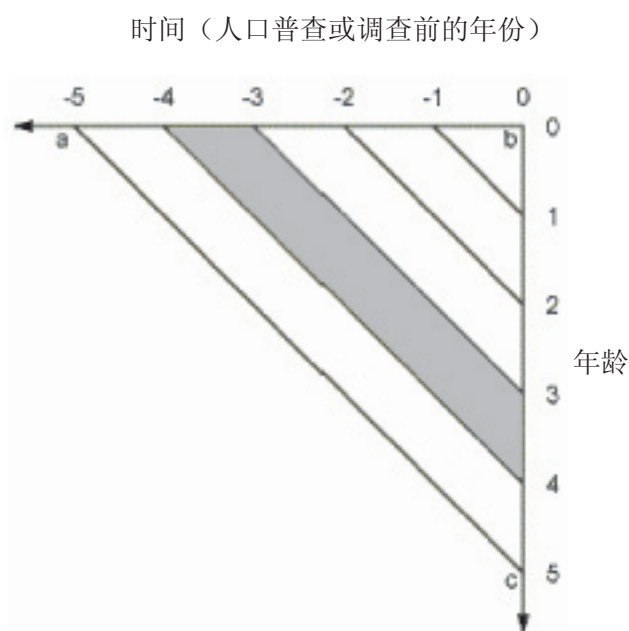
711. 承受风险原则显示了一种计算婴儿死亡率的不同方法：用某一特定年份出生的儿童数量去除这些儿童中过第一个生日以前的死亡数，也就是说，用直线  $ab$  代表的儿童人数去除平行四边形  $abed$  代表的这些儿童中的死亡数。

712. 传统计算方法的合理性是双重的。首先，如果出生数和婴儿死亡数在各年之间的变化不是很快的话，正方形  $abcd$  中的死亡数将与平行四边形  $abed$  中的死亡数大体相当，因为与三角形  $adc$  相应的死亡数将大体等于三角形  $bed$  代表的死亡数。其次，某一特定年份发生的婴儿死亡数通常比平行四边形  $abed$  代表的死亡数更容易从标准死亡登记表中获得。

#### F. 存活倒推法

713. 根据人口普查和调查的年龄分布对出生数和死亡率进行估计的存活倒推法见第五章A节。

图A.6. 存活倒推法\*



\*见第五章A节。

图A6展示了这种方法的几个方面。在单一基准时间的人口普查和调查中，时间轴将代表日历时间，其中时间0代表基准时间。如果访问是在一段时间内进行的，而且收集的信息是访问时的信息，在这样的调查中，时间轴将代表访问前的年份。为简单说明起见，这里假设根据人口普查进行估计。

714. 从时间0向下延伸的纵（年龄）轴代表人口普查或调查时人口中的个人。向左延伸的横（时间）轴代表人口普查前年份的人口中的出生数。

715. 年龄轴线段代表人口普查时特定年龄组中的个人。例如，实足年龄3岁和实足年龄4岁之间的线段代表周岁为3岁的个人。时间轴线段代表人口普查前各年出生的人群，这样，连接时间-3和-4的部分代表人口普查前第四年的出生数。

716. 斜线和它们形成的梯形表明了人口普查前各年出生的儿童、人口普查时这些儿童中的存活者以及人口普查前死亡的儿童之间的关系。例如，人口普查时周岁为3岁的儿童是人口普查前第四年出生的儿童中的存活者。人口普查前发生的该组儿童的死亡以连接时间轴和年龄轴上相应间隔的阴影梯形表示。

717. 存活倒推法基于任何同期群中的出生数相当于人口普查时的存活数加上该组群中人口普查基准时间以前的死亡数这一简单观察数据。根据人口普查的年龄分布可以知道人数（当然可能有各种误差）。死亡数根据对婴儿和儿童死亡率水平的了解进行估计。

718. 图A6说明了分各年的年龄组和时间期间的存活倒推法。这种方法可能也适用于综合年龄组。例如，直线bc代表的人口普查时周岁为0至4岁的人是直线ab代表的人口普查前五年间出生的人当中的存活者。出生数等于存活数加上死亡数，以三角形abc表示。

### G. 亲生子女法

719. 第五章B节介绍了生育率估计的亲生子女法。这种方法包含若干技术性调整，以将亲生子女列表中的数据转换成标准的分年龄组出生率。这些调整的性质可通过适当绘制的列克西斯图来说

明。如上一节所述，假设数据来自人口普查可以简化说明。

720. 表3所示的亲生子女列表的每一行都给出了按儿童年龄与其母亲配对的儿童人数。例如，年龄为25岁的一行中的数字表明了25岁妇女的亲生子女按普查基准时间的年龄分布。周岁为0岁的儿童出生于人口普查前的年份，以图A7中的平行四边形*behd*表示，该平行四边形是时间轴上-1和0之间的间隔下面的垂直矩形区域与从年龄轴上实足年龄25岁和26岁之间的间隔向上和向左延伸的一组斜线交叉形成的。

721. 图A7是从亲生子女表3选出的一些组格。具体来说，交叉方法表明：

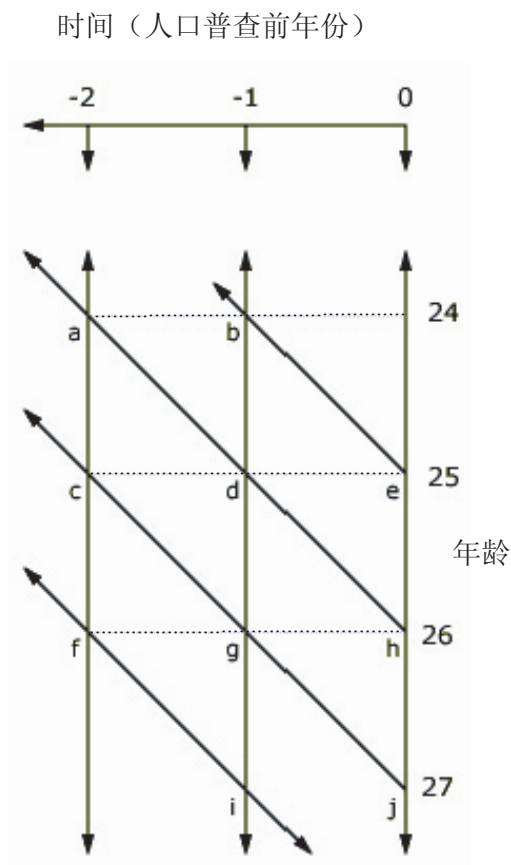
- 平行四边形*behd*代表人口普查时年龄为25岁的妇女在普查前一年所生的亲生子女
- 平行四边形*dhjg*代表人口普查时年龄为26岁的妇女在普查前一年所生的亲生子女
- 平行四边形*adgc*代表人口普查时年龄为26岁的妇女在普查前第二年所生的亲生子女
- 平行四边形*cgif*代表人口普查时年龄为27岁的妇女在普查前第二年所生的亲生子女

其中妇女的年龄是周岁。相应的出生数是乘以存活倒推和非亲生子女系数得出的。不过，出生数并不是标准的分年龄组出生率的分子。例如，人口普查前第二年25岁年龄组妇女的出生率分子是图A7中正方形*cdgf*代表的一组数字。从理论上来说，从亲生子女列表获得这些出生数的方法包括以下步骤：

- 用适当计算出的分离系数乘以平行四边形*adgc*代表的出生数来计算三角形*cdg*代表的出生数
- 以同样的方法从平行四边形*cgif*代表的出生数计算出三角形*cgf*代表的出生数
- 将这两个三角形代表的出生数相加得到正方形*cdgf*代表的出生数

722. 在实践中，通常假定分离系数为1/2。分离系数实际上可能与1/2相差很多，但有时会取消一些误差，剩下的误差与其他误差相比通常很小，从而保证使用更精确的分离系数。若应用于质量很高

图A. 7. 亲生子女法\*



\*见第五章B节。

的数据，在估计分年龄组出生率的分子时可能需要更复杂的程序。

#### H. 平均生育数

723. 图A8是人口普查或调查时选定年龄组妇女平均生育数的列克西斯图。这里假设所有的生育都是15至49岁妇女的生育。

724. 更年轻的妇女，如人口普查或调查时20至24岁妇女的平均生育数（阴影区1）反映了生育时间和水平。仍处在育龄期的年龄较大的妇女，如35至39岁妇女的平均生育数（阴影区2），是育龄总生育率的更好指数。50岁以上妇女的平均生育数大体相当于相应的同期出生群的育龄总生育率（阴影区3和4）。数据提供的只是一个近似值，因为它必定将出生数从人口普查以前死亡的妇女中扣除。

725. 某一人口中任何同期出生群妇女的平均

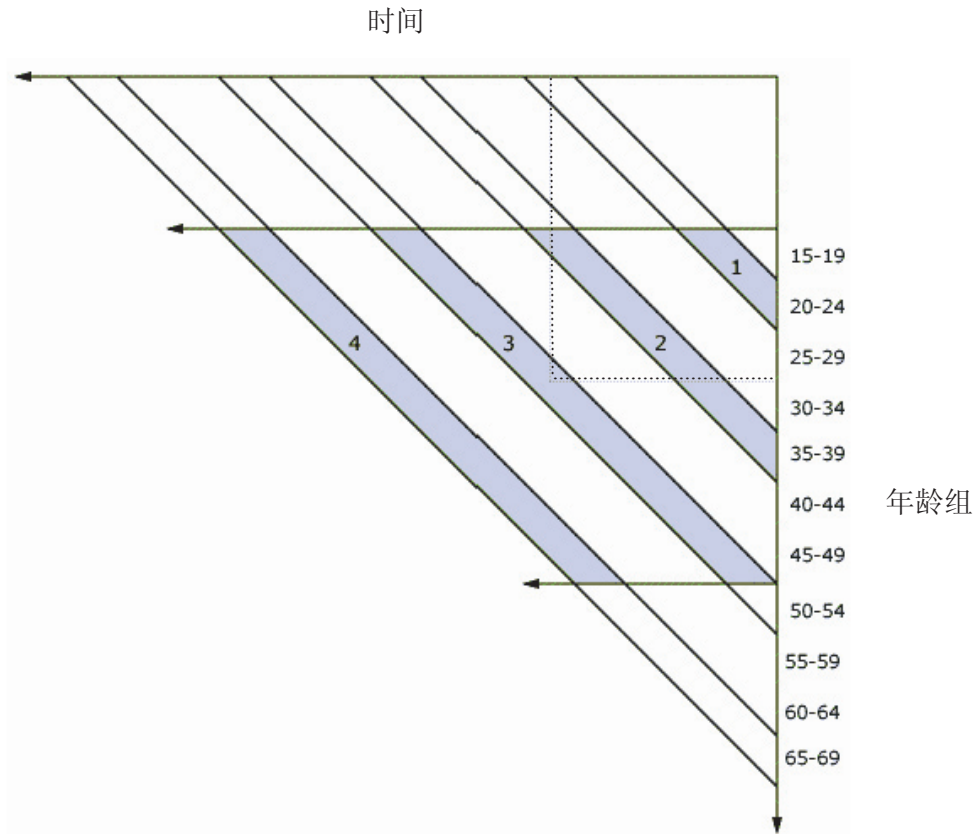
生育数大约相当于同期出生群成员达到该组群平均生育年龄时人口的总和生育率的近似值。图A8说明了人口普查时50至54岁妇女的平均生育数。这些妇女的生育数以阴影区3表示。生育数除以妇女人数得出的就是该组群的平均生育数。

726. 平均生育年龄变化不大，在数据缺乏的情况下可以定为30岁左右。阴影区3代表的同期出生群成员在人口普查或调查日前的20年平均达到了30足岁。30岁的水平虚线代表生育的平均年龄。垂直虚线表明平均生育数大体相当于人口中总和生育率的时间。

#### I. 生育史

727. 第五章F节介绍的生育史是收集生育率和婴儿与儿童死亡率数据的一个重要方法。其中采用的非常详细的问题为调查员提供了广泛考察的机

图A. 8. 平均生育数



会。可以对此进行探索，以便从被调查者获得其他方法不可能获得的更加精确的数据。

728. 不过，生育史数据受到重要的技术局限性的限制，特别是在（情况通常如此）只收集到育龄妇女的生育史时。排除50岁以上的妇女意味着收集到的数据只是“不完整组群”，即调查以后还将生育孩子的妇女组群。

729. 15至49岁妇女的生育史提供了调查访谈前35年的出生数据，但这些数据因为调查时被取消而逐渐变得不完整。图A9说明了这一点。下面的分析为简单起见，假定所有调查访谈都是在同一时点发生的。

730. 调查前35年的出生总数以正方形 $abcd$ 表示。在0这一时点上15至49岁妇女的生育史只提供了三角形 $abd$ 代表的生育信息。三角形 $adc$ 代表的生育

数是50岁以上妇女在0时的生育数，这方面的信息没有收集到。

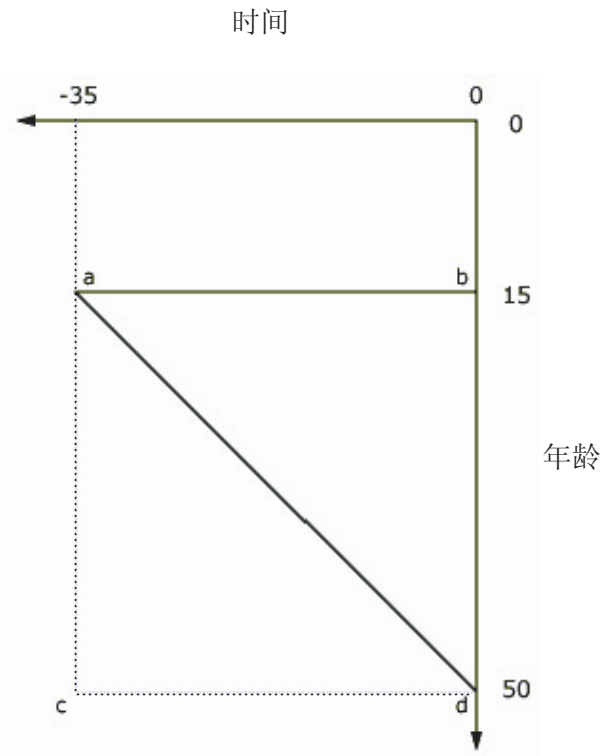
731. 从图中可以明显看出，尽管生育史调查提供了调查前35年的生育信息，但年份越早，信息越不完全。

732. 严格说来，调查数据甚至不足以捕获调查前一年的总和生育率，因为这一年年龄为49周岁的妇女的生育数只在她们于当年达到49足岁时才得以表示。年初为49周岁的母亲在调查时的年龄为50岁，因此未被包含在内。

733. 不过，很少有妇女在育龄期末生育，育龄妇女的生育史提供了令人满意的调查前最初几年的总和生育率近似值。然而，就更早的年份而言，总和生育率只能通过使用一些对三角形 $adc$ 代表的生育数取近似值的方法获得。



图A. 9. 生育史数据



## 术 语 表

注：参考资料酌情注明了章、节和小节。因此(I)指第一章，(II.B)指第二章B节，(III.C.3)指第三章C节第3小节。

### 年 龄

实足年龄指任何人在任何给定时间自出生以来的时间流逝。周岁，有时也称为最近生日时的年龄，是小于实足年龄的最大整数。“年龄”本身可根据上下文指实足年龄或周岁，它几乎总是指这两个年龄中的一个。

### 周 岁

见“年龄”。

### 平衡方程式

见“人口方程式”。

### 出 生

见“活产”。

### 生 育 史

某个妇女的所有活产数，或某一特定日期以前的所有出生数，提供出生儿童的出生日期和性别。如果生育史包括某一特定日期以前的出生数，该清单通常还包括截至当日出生儿童是否存活的信息，如果未存活的话，包括其死亡日期和/或死亡年龄。生育史有时是部分生育史，仅限于某一特定期间如某次调查前五年的出生数，或某一特定胎次的出生数，例如按1-4的胎次或最近的出生。

### 生育史调查

包括生育史问题在内的住户调查。

### 死 因

指导致或促成死亡的所有疾病、病状或伤害及造成任何此类伤害的事故或暴力的情况。就生命统计目的而言，诸如心跳衰竭或衰弱等死亡症状或方式不认为是死因。又见“基本死因”和“辅助死因”。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

### （死因）证明人

指经法律授权以规定格式出具证明的人，说明死亡的基本原因和辅助原因以及与死亡事件有关的其他事实，并将其提交地方登记官或其他适当人士。证明通常由在死者患病后期护理他/她的医生出具，或者，如果死者患病后期无医生护理，或死亡系由暴力或伤害所致，则由法医官员（如验尸官或体检医生）出具。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

### 民事登记

根据各国的法律要求，对生命事件（活产、死亡、死胎、结婚和离婚）和法令、法律与法规规定的其他与人口有关的民事事件的发生和特点进行的连续、永久、强制和普遍性记录。它建立并提供关于此类事件的法律文件。民事登记记录也是生命统计的最佳资料来源。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

### 辅助死因

促成死亡但与直接引起死亡的疾病或病状无关的一个重要病状。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

### 数 据

1. 关于某些定义明确的集合所包含的实体性质的系统信息，如人口普查或调查产生的个人记录或民事登记系统产生的出生或死亡记录。此类数据可称为“微观”或“单位记录”或“个人一级”数据。这种意义上的数据与数据集同义。虽然记录中包含的信息可能是定量信息，集合的定义应该是原文的定义，以便数据始终包含某一定性因素。

2. 这些数据产生的数字信息，如从人口普查数据得出的各种年龄-性别组的人数表。这种数据可称为“宏观”或“集合”或“表格”数据。在统计领域的术语中，称为统计量。

3. 一般定量信息，包括估计、指标和各种统计资料。

## 数据集

说明某些定义明确的统计集合中所包含实体的一组记录，如某一特定年份某个国家的所有出生数，或某一特定时点某个国家人口中的所有人数。在最简单的情况下，每个实体由提供关于这一实体的相关信息的记录表示。一般来说，为每个实体提供的信息是相同的，例如，要么提供每个人的婚姻状况，要么谁的也不提供。不过，不适用的信息必须除外。例如，第一次结婚时的年龄只对结过婚的人有意义。“数据集”的字面意思与“数据”的第一种含义相同，但数据集是一种适于在电脑上处理的形式，并且能够提供关于记录格式和编码的辅助信息。

## 死亡

死亡是指在活产后的任何时间所有生命迹象的永久消失（出生后生命功能的停止，而且不能复活）。这一定义不包括死胎（见死胎）。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

## 人口方程式

说明任何时期间人口规模的变化是这一时期增加的人口数量和该时期减少的人口数量之间差额的同义反复。

## 估计

1. 在通常的用法中，用作动词是指得出某些不能直接观察或测量的数量值的行为；用作名词是指这一估计行为的结果。2. 在统计领域，用于估计统计模型中某个参数的统计量。

## 事件登记

每当某一规定类型的事件发生时就进行更新的数据集，如某一定义明确的人口发生的出生或死亡登记记录。又见“人口登记”。

## 实足年龄

见“年龄”。

## 承受风险原则

（第四章B节第1小节）在计算人口事件的发生率时，分子中的事件和分母中的个人必须满足以下条件的原则：分子中的所有事件必须是发生在分母中个人身上的事件，发生在分母中个人身上的所有事件必须计算在分子中。又见“度过的人年”。

## 死胎

妊娠产物在完全从母亲体内产出或取出之前的死亡，不管妊娠期的长短；该死亡由以下事实表明，即胎儿从母体分离后没有呼吸或没有任何其他生命迹象，如心脏跳动、脐带震动或肌肉明显的自主运动。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

## 估算

用统计上似乎合理但不一定正确的估算值代替某项缺失或不能接受的报告值的方法。估算也可用于解决两个或两个以上数值之间逻辑上的矛盾，例如用与另一个报告值相一致的估算值来代替某个报告值。

## 申报人

（民事登记）根据法律规定负责向地方登记官报告某一生命事件发生的事实并提供与该事件有关的所有信息和特征的个人。以这一报告为基础，该事件由地方登记官合法进行登记。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

## 活产

活产是妊娠产物从母亲体内完全产出或取出的结果，不管妊娠期的长短，该产物从母体分离后能够呼吸或表现出任何其他生命迹象，如心跳、脐带的震动或肌肉明显的自主运动，无论脐带是否已被剪断或胎盘是否缠附；这样出生的每个产物被视为活产。（术语表，《关于人口动态统计系统的原则和建议》，联合国，2001年b。）

## 平均生育年龄

分年龄组出生率年龄时间表的平均年龄。分年龄组出生率可指同期出生群或某时期。

## 产妇死亡

产妇死亡指妇女在妊娠期间或终止妊娠42天以内死于与妊娠或妊娠处理有关的或因妊娠或妊娠处理而恶化的任何原因，而不是偶然或意外原因导致的死亡——不管妊娠期和地点。（术语表，《关于生命统计系统的原则和建议》，联合国，2001年b。）

## 产妇死亡率

以某一特定时间内（通常为一个日历年）由于孕产原因而发生的相对于活产数的死亡数为基础的

生命统计比率，即在特定的某一年中发生在某一特定地理区域内每100 000(或10 000)活产中死于妊娠、分娩和产褥期的直接产科并发症, 手术、疏忽或不当处理或其后果, 或先前存在的疾病或妊娠期间感染的疾病所导致的间接产科原因以及并非死于直接产科原因但死于因妊娠的生理影响而恶化的原因的妇女数。(术语表, 《关于生命统计系统的原则和建议》, 联合国, 2001年b。)

## 胎 数

妇女在任何特定时间的生育数是指她截至某一特定时间的活产数。

## 生育累进率

至少已有*i*次活产的一组妇女中继续发生第*i*+1次活产的妇女所占的比例。*i*>0是从第*i*次到第*i*+1次生产的累进率, *i*=0是首次生产的累进率。

## 亲生子女法

根据人口普查或户口调查估计分年龄组出生率的方法。该方法提供数据收集工作的基准时间以前15年或15年以上的出生率。它包括对住户中个人的普查或调查记录进行处理, 以确定年龄在15岁以下的每个人的母亲目前是否存在于住户中, 如果是的话, 母亲是哪个人。详情见第五章B节。

## 度过的人年

在某一特定期限的某个人口中度过了*x*年的人给这一期间的人口增加了*x*个活过的人年。任何期间某一人口中度过的人年是这一人口的所有成员在这一期间的任何时间活过的人年之和。任何期间某一人口中活过的人年也可以说成是人口规模与特定期限坐标图下面的区域。某一特定期限某一人口中活过的人年大约相当于人口规模初期和末期的平均期间长度。任何分组人口的人年完全可以将分组人口视为一个人口来定义。

## 人 口

“人口”在人口统计学和统计学中都是一个定义概念, 但在两个领域表示的概念非常不同。在两种情况下它指的都是一个集合。不过, 在人口统计学中, 该集合随时间而变化, 因为新成员不断加入, 现有成员不断离开, 而在统计学中, 它是静止的。1. 人口统计学: 其成员随时间而变化的一组人, 在这组人中, 成员经常首先从地理位置来

定义, 但也可以从各种个人特征来定义, 如“女性人口”或“使用避孕用具的人口”。2. 统计学: .....任何有关客体的集合, 无论是有生命的, 还是无生命的.....(Stuart和Ord, 1987年, 第1卷, 第1.1-1.3节)。

## 人口登记

说明组成某一人口的成员和个人某些特征的数据集。数据集中的记录随着人口成员的变化而更新, 记录的内容则随着它们说明的个人特征的变化而更新。这两种更新均可视为由某些事件的发生而引起。在第一种情况下, 事件是指改变人口成员的事件, 是人口中成员的增加或死亡。在第二种情况下, 事件指是人口中个人特征的变化。又见“事件登记”。

## 比 率

在人口统计研究中以多种方式使用的一个术语, 但最适于指某一特定期限发生的事件除以这一期间的平均人口(或适当的分组人口)所得的数字。该词也可用于指分组人口和所有人口之间的比率: 如入学率或识字率(Wilson, 1985年)。

## 登 记

用作动词, 指记录关于最近发生的某一事件信息的行为。用作名词, 见“事件登记”和“人口登记”。

## 样 本

为统计目的用于说明一组人或一组事件整个组别特征的分组。

## 统 计 量

1. 在通常的用法中, 指说明事实某个方面的数字。2. 在统计领域, 指可从特定数据计算的任何数量(统计量是观察资料的一个功能)。

## 制 表

用作动词, 指将数据集中的每项记录进行处理以确定若干类别中每个类别的记录总数; 用作名词, 指这种制表产生的数据结构。

## 总和生育率

生育水平的一个基本指标, 通过计算所有育龄的年龄别出生率总和而得出。它可以解释为在育龄期末仍然存活的妇女如果经历年龄别出生率在一生

中将拥有的预期子女数。

### **基本死因**

引起直接导致死亡的一系列事件的疾病或伤害,或造成致命伤害的事故或暴力情况。基本死因被用作死亡统计制表的基础。又见“辅助死因”。(术语表,《关于生命统计系统的原则和建议》,联合国,2001年b。)

### **生命事件**

(民事登记)活产、死亡、死胎、结婚、离婚、收养、合法化、确认父母身份、婚姻失效或合

法分居的发生。(术语表,《关于生命统计系统的原则和建议》,联合国,2001年b。)

### **生命事件记录**

(民事登记)在民事登记簿中登记的法律文件,用于证明某一生命事件的发生和特征。(术语表,《关于生命统计系统的原则和建议》,联合国,2001年b。)

### **加权样本**

具有某些特征的个人或事件有不同列入机会的样本。

## 参考资料

- Banister, Judith (1979). Census questions on fertility and child mortality: problems with questionnaire design. *Asian and Pacific Census Forum* vol. 6, No.1.
- Bennett, Neil G., and Shiro Horiuchi (1981). Estimating the completeness of death registration in a closed population. *Population Index* vol. 47, No.2.
- \_\_\_\_\_ (1984). Mortality estimation from registered deaths in developed countries. *Demography* vol. 21, No.2.
- Blacker, J. G. C., and J. Mukiza-Gapere (1988). The indirect measurement of adult mortality in Africa. In *African Population Conference, Dakar, 1988*. Liège, Belgium: International Union for the Scientific Study of Population (IUSSP).
- Bongaarts, John, and Robert G. Potter (1983). *Fertility, Biology, and Behavior: An Analysis of the Proximate Determinants*. New York: Academic Press.
- Brass, William (1979). Evaluation of birth and death registration using age distributions and child survivorship data. *Asian and Pacific Census Forum* vol. 5, No.3.
- \_\_\_\_\_ (1985). A simple approximation for the time-location of estimates of child mortality from proportions dead by age of mother. In *Advances in Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Occasional Publication, Centre for Population Studies. London: London School of Hygiene and Tropical Medicine.
- Brass, William, and Sheila Macrae (1984). Childhood mortality estimated from reports on previous births given by mothers at the time of a maternity I. Preceding birth technique. *Asian and Pacific Census Forum* vol. II, No.2.
- \_\_\_\_\_ (1985). Childhood mortality estimated from reports on previous births given by mothers at the time of a maternity II. Adapted multiplier technique. *Asian and Pacific Census Forum* vol. 11, No.4.
- Casley, D. J and D. A. Lury (1981). *Data Collection in Developing Countries*. Oxford, United Kingdom Clarendon Press.
- Chackiel, J., and H. Orellana (1985). Adult female mortality trends from retrospective questions about maternal orphanhood included in censuses and surveys. In *International Population Conference, Florence, 1985*, vol. 4, Liège, Belgium: International Union for the Scientific Study of Population.
- China Population Information Centre (1984). *Analysis of China's National One-per-Thousand-Population Fertility Sampling Survey*. Beijing: China Population Information Centre.
- Cho, Lee-Jay (1976). *Introduction to Censuses of Asia and the Pacific: 1970-74*. Honolulu: East-West Center.
- Cho, Lee-Jay, and Robert L. Hearn (1984). *Censuses of Asia and the Pacific: 1980 Round*. Honolulu: East-West Center.
- Cho, Lee-Jay, Robert D. Retherford and Minja Kim Choe (1986). *The Own-Children Method of Fertility Estimation*. Honolulu: East-West Center.
- Cleland, John (1996). Demographic data collection in less developed countries. *Population Studies* vol. 50, No 3.
- Cleland, John, and Chris Scott, eds. (1987). *World Fertility Survey: An Assessment*. Oxford, UK: Oxford University Press.
- Coale, Ansley J. (1984). Construction of a life table from accurate enumeration of a closed population in two censuses. *Population Index* vol. 50, No.2.
- Coale, Ansley J., Lee-Jay Cho and Noreen Goldman (1980). *Estimation of Recent Trends in Fertility and Mortality in the Republic of Korea*. Washington, D.C.: National Academy of Sciences.
- Dechter, Aimée R., and Samuel H. Preston (1991). Age misreporting and its effects on adult mortality estimates in Latin America. *United Nations Population Bulletin Nos.31/32*.
- Del Tufo, M. V. (1949). *A Report on the 1947 Census of Population*. Kuala Lumpur: The Government Printer, Federation of Malaya.
- El Badry, M. R. (1961). Failure of enumerators to make entries of zero: errors in recording childless cases in population censuses. *Journal of the American Statistical Association*, vol. 56, No.296.
- Ewbank, D. C. (1981). *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns and Consequences for Demographic Analysis*. Committee on Population and Demography, Report No. 4.

- Washington, D.C.: National Academy Press.
- Feeney, Griffith (1976). Tabulation of census and survey data on child survivorship. *Asian and Pacific Census Newsletter*, vol. 3, No. 1.
- \_\_\_\_\_ (1983). Population dynamics based on birth intervals and parity progression. *Population Studies*, vol. 37, No.1.
- \_\_\_\_\_ (1988). The use of parity progression models in evaluating family planning programmes. In *African Population Conference: Dakar 1988*, vol. 3, pp. 7.1.17-17.1.30. Liège, Belgium: International Union for the Scientific Study of Population.
- \_\_\_\_\_ (1991). Child survivorship estimation: methods and data analysis. *Asian and Pacific Population Forum*, vol. 5, Nos. 2-3.
- Feeney, Griffith, and Feng Wang (1993). Parity progression and birth intervals in China. *Population and Development Review*, vol. 19, No.1.
- Feeney, Griffith, and John A. Ross (1984). Analyzing open birth interval distributions. *Population Studies* vol. 38, No.3.
- Feeney, Griffith, and Yasuhiko Saito (1985). Progression to first marriage in Japan: 1870-1980. NUPRI Research Paper No. 24, Tokyo: Nihon University, Population Research Institute.
- Feeney, Griffith, and Jingyuan Yu (1987). Period parity progression ratio measures of fertility in China. *Population Studies*, vol. 41, No.3.
- Goldfarb, Charles F. (1990). *The SGML Handbook*. Oxford, UK: Clarendon Press.
- Hanenberg, Robert (1991). *The Computation and Analysis of Vital Rates from Demographic Surveys: Using Standard Microcomputer Software*. Bangkok: Economic and Social Commission for Asia and the Pacific.
- Henry, Louis (1980). *Fertility of Marriages: A New Method of Measurement*. Population Studies Translation Series, No. 3, Economic and Social Commission for Asia and the Pacific. New York: United Nations. Originally published in 1953 as *Fécondité des Mariages: Nouvelle Méthode de Mesure*.
- Hill, K. (1987). Estimating census and death registration completeness. *Asian and Pacific Census Forum*, vol. 1, No.3.
- Hill, K., and J. Trussell (1977). Further developments in indirect mortality estimation. *Population Studies*, vol. 21, No.2.
- Hull, Terence H. (1995). Levels and trends of infant and child mortality in Indonesia, estimated in 1978: a methodological review. In *Kecendrungan dan Faktor Penentu Fertilitas dan Mortalitas di Indonesia [Trends and Determinants of Fertility and Mortality in Indonesia]*, Aris Ananta, ed. Jakarta: Kantor Menteri Negara Kependudukan/BKKBN.
- Hull, Terence H., and Sunaryo (1978). *Levels and Trends of Infant and Child Mortality in Indonesia*. Working Paper 15. Yogyakarta, Indonesia: Universitas Gadjah Mada, Population Institute.
- Institute for Resource Development/Macro International. (1991). *Demographic and Health Surveys World Conference, August 5-7, 1991, Washington, D.C.* Columbia, Maryland: Macro International, Inc.
- Alan Stuart and J. Keith Ord (1987). *Kendall's Advanced Theory of Statistics*, vols. 1 and 2. New York: Oxford University Press.
- Luther, Norman Y., and Lee-Jay Cho (1988). Reconstruction of birth histories from census and household survey data. *Population Studies*, vol. 42, No.3.
- Luther, Norman Y., Neramit Dhanasakdi and Fred Arnold (1986). *Consistent correction of census and vital registration data for Thailand, 1960-1980*. Papers of the East-West Population Institute, No. 103. Honolulu: East-West Center.
- Luther, Norman Y., Griffith Feeney and Weimin Zhang (1990). One-child families or a baby boom? Evidence from China's 1987 one-per-hundred survey. *Population Studies*, vol. 44, No.2.
- Macura, Miroslav (1972). Estimates of the completeness of registration of births and infant deaths in Yugoslavia and its main provinces from the late 1940s to 1961. Doctoral dissertation, Princeton University.
- Marks, Eli S., William Seltzer and Karol J. Krotki (1974). *Population Growth Estimation: A Handbook of Vital Statistics Estimation*. New York: The Population Council.
- Myanmar Immigration and Population Department (1995). *Population Changes and Fertility Survey 1991*. Yangon: Ministry of Immigration and Population.
- Myanmar Department of Population (1998). *Myanmar Fertility and Reproductive Health Survey 1997, Preliminary Report* Yangon: Ministry of Immigration and Population.
- National Academy of Sciences (1981). *Collecting Data for the Estimation of Fertility and Mortality*. Report No. 6, Committee on Population and Demography. Washington, D.C.: National Academy Press.
- Preston, Samuel (1980). Estimating adult female mortality from the number of children surviving. *Asian and Pacific Census Forum*, vol. 6, No. 4.

- \_\_\_\_\_ (1985). Mortality in childhood: lessons from the WFS. In *Reproductive Change in Developing Countries: Insights from the World Fertility Survey* J.G. Cleland and John Hobcraft, eds. London: Oxford University Press.
- Retherford, Robert D., and Iqbal Alam (1985). *Comparison of fertility trends estimated alternatively from birth histories and own children*. Papers of the East-West Population Institute No. 94. Honolulu: East-West Center.
- Retherford, Robert D., and others (1987). Fertility trends in Pakistan: The decline that wasn't. *Asian and Pacific Population Forum*, vol.1.No.2.
- Retherford, Robert D., and G. Mujtaba Mirza (1982). Evidence of age exaggeration in demographic estimates for Pakistan. *Population Studies*, vol.36, No.2.
- Retherford, Robert D., Vinod K. Mishra and G. Prakasam (2000). *How Much Has Fertility Declined in Uttar Pradesh?* National Family Health Survey Subject Reports, No. 17. Mumbai, India: International Institute for Population Sciences; and Honolulu:East-West Center Population and Health Studies.
- Rindfus, Ronald R., and James A. Sweet (1977). *Postwar Fertility Trends and Differentials in the United States*. New York: Academic Press.
- Rutstein, Shea O. (1984). Infant and child mortality: levels, trends and demographic differentials. revised edition. *WFS Comparative Studies No. 43*. Voorburg, Netherlands: International Statistical Institute.
- Saw Swee-Hock (1964). A note on the under-registration of births in Malaya during the intercensal period 1947-1957. *Population Studies*, vol. 18, No.1(July 1964).
- Stupp, Paul W. (1988). Estimating intercensal age schedules by intracohort interpolation. *Population Index*, vol. 54, No. 2.
- Sullivan, Jeremiah M., Shea Oscar Rutstein and George T. Bicego (1994). Infant and child mortality. *Demographic and Health Surveys Comparative Studies No. 15*. Calverton, Maryland: Macro International, Inc.
- Timaeus, I. M., Chackiel, J. and Ruzicka, L., eds. (1996). *Adult Mortality in Latin America*. Oxford: Clarendon Press.
- The Unicode Consortium (2000). *The Unicode Standard, Version 3.0*. Mountain View, California. Available from <http://www.unicode.org>.
- United Nations (1969). *Methodology and Evaluation of Population Registers and Similar Systems*. Studies in Methods, No. 15. Sales No. E.69.XVII.15.
- \_\_\_\_\_ (1973). *Principles and Recommendations for a Vital Statistics System, Revision 1*. Sales No. E.73.XVII.9.
- \_\_\_\_\_ (1982). *Non-sampling Errors in Household Surveys: Sources, Assessment and Control - Preliminary Version*. National Household Survey Capability Programme Technical Study DP/UN/INT-81-041-2.
- \_\_\_\_\_ (1983). *Manual X: Indirect Techniques for Demographic Estimation*. Population Studies, No. 81. Sales No. E.83.XIII.2.
- \_\_\_\_\_ (1984). *Handbook of Household Surveys. Studies in Methods, Series F, No. 31*.Sales No. E.83.XVII.13.
- \_\_\_\_\_ (1985). *Development and Design of Survey Questionnaires*. National Household Capability Programme. INT-84-014.
- \_\_\_\_\_ (1986). *Sampling Frame and Sample Designs for Integrated Household Survey Programs - Preliminary Version*. National Household Survey Capability Programme. DP/UN/INT-84-014/5E.
- \_\_\_\_\_ (1993). *Sampling Errors in Household Surveys*. National Household Capability Programme. INT-92-980-15E.
- \_\_\_\_\_ (1998a). *Principles and Recommendations for Population and Housing Censuses, Revision 1*. Sales No. E.98.XVII.8.
- \_\_\_\_\_ (1998b). *Handbook on Civil Registration and Vital Statistics Systems: Management, Operation and Maintenance*. Sales No. E.98.XXVII.11.
- \_\_\_\_\_ (1998c). *Handbook on Civil Registration and Vital Statistics Systems: Preparation of a Legal Framework*. Sales No. E.98.XXVII.7.
- \_\_\_\_\_ (1998d). *Handbook on Civil Registration and Vital Statistics Systems: Developing Information, Education and Communication*. Sales No. E.98.XXVII.4.
- \_\_\_\_\_ (1998e). *Handbook on Civil Registration and Vital Statistics Systems: Computerization*. Sales No. E.98.XXVII.10.
- \_\_\_\_\_ (1998f). *Handbook on Civil Registration and Vital Statistics Systems: Policies and Protocols for the Release and Archiving of Individual Records*. Sales No. E.98.XVII.6.
- \_\_\_\_\_ (2000a). *Handbook on Census Management for Population and Housing Censuses*. Sales No. E.00.XVII.15.
- \_\_\_\_\_ (2000b). *Handbook on Geographic Information Systems and Digital Mapping*. Sales No.



E.00.XVII.12.

\_\_\_\_\_ (2001a). *Handbook on Population and Housing Census Editing*. Sales No. E.00.XVII.9.

\_\_\_\_\_ (2001b). *Principles and Recommendations for a Vital Statistics System, Revision 2*. Sales No. E.01.XVII.10.

Verma, Vijay (1980). *Basic Fertility Measures from Retrospective Birth Histories*. World Fertility Survey Technical Bulletins, No. 4/TECH. 1407. The Hague: International Statistical Institute.

Vincent, Paul (1946). L'Utilisation des statistiques des familles. *Population*, vol. 1.

Whelpton, Pascal K. (1954). *Cohort Fertility: Native White Women in the United States*. Princeton, New Jersey: Princeton University Press.

Willcox, Walter F. (1940). *Studies in American Demography*. New York: Russell and Russell.

World Health Organization(1993) International Statistical Classification of Diseases and Related Health Problems, Tenth Revision vol. 2, Instruction Manual. Geneva.





---

### كيفية الحصول على منشورات الأمم المتحدة

يمكن الحصول على منشورات الأمم المتحدة من المكتبات ودور التوزيع في جميع أنحاء العالم . استعلم عنها من المكتبة التي تتعامل معها أو اكتب إلى : الأمم المتحدة ، قسم البيع في نيويورك أو في جنيف .

#### 如何购取联合国出版物

联合国出版物在全世界各地的书店和经售处均有发售。请向书店询问或写信到纽约或日内瓦的联合国销售组。

#### HOW TO OBTAIN UNITED NATIONS PUBLICATIONS

United Nations publications may be obtained from bookstores and distributors throughout the world. Consult your bookstore or write to: United Nations, Sales Section, New York or Geneva.

#### COMMENT SE PROCURER LES PUBLICATIONS DES NATIONS UNIES

Les publications des Nations Unies sont en vente dans les librairies et les agences dépositaires du monde entier. Informez-vous auprès de votre libraire ou adressez-vous à : Nations Unies, Section des ventes, New York ou Genève.

#### КАК ПОЛУЧИТЬ ИЗДАНИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ

Издания Организации Объединенных Наций можно купить в книжных магазинах и агентствах во всех районах мира. Наводите справки об изданиях в вашем книжном магазине или пишите по адресу: Организация Объединенных Наций, Секция по продаже изданий, Нью-Йорк или Женева.

#### COMO CONSEGUIR PUBLICACIONES DE LAS NACIONES UNIDAS

Las publicaciones de las Naciones Unidas están en venta en librerías y casas distribuidoras en todas partes del mundo. Consulte a su librero o diríjase a: Naciones Unidas, Sección de Ventas, Nueva York o Ginebra.

---