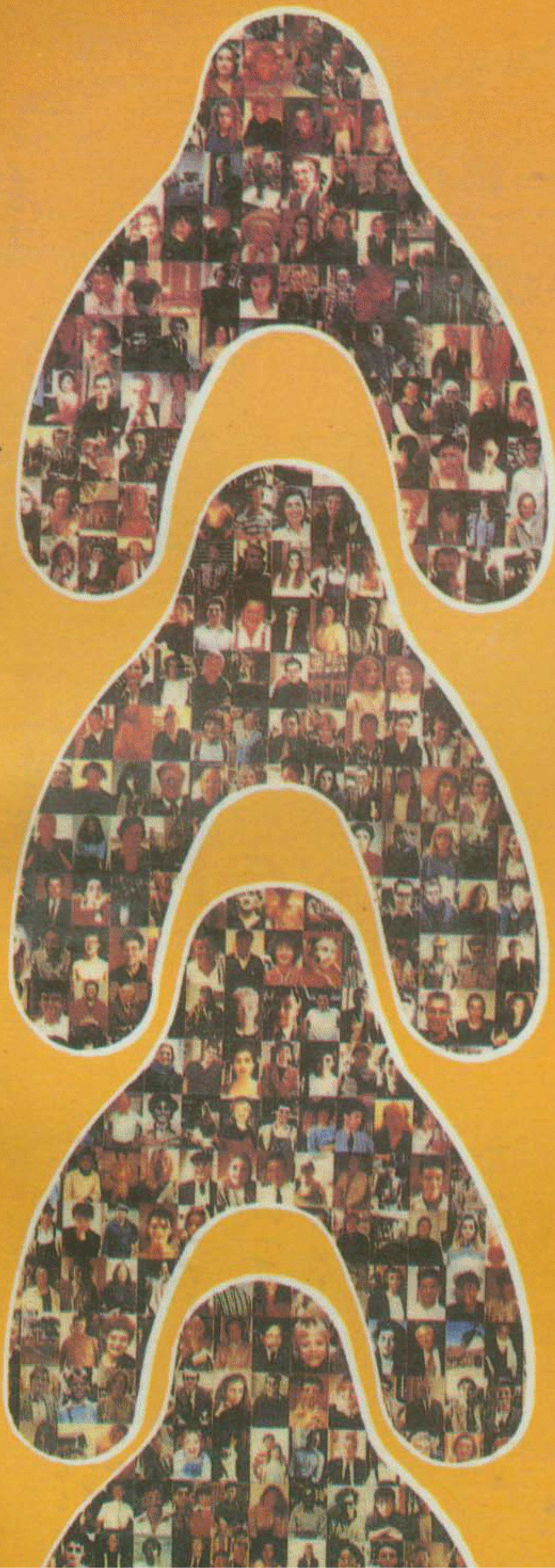


人口分析方法与应用

曾毅 编著

北京大学出版社



《人口分析方法与应用》一书把欧、

美与中国人口分析方法和应用溶为一体，全

面系统地阐述了人口分析的基础知识、基本概念、量测指

标、人口模型及研究方法。它是人口学及相关专业的必备教材或参考书。

人口分析方法与应用

曾毅 编著

北京大学出版社

目 录

绪 言	(1)
第一章 队列分析的基本原理	(1)
1.1 基本概念	(1)
1.1.1 队列的定义	(1)
1.1.2 队列的参照时间与事件的时间量度	(1)
1.1.3 列克西斯图	(2)
1.1.4 应用与练习	(4)
1.2 不考虑其他人口事件影响的队列分析	(5)
1.2.1 平均强度、平均年龄(或期间)与中位数	(6)
1.2.2 年龄(期间)别概率	(8)
1.2.3 年龄(期间)别发生/风险率	(9)
1.2.4 发生/风险率与概率之间的关系	(11)
1.2.5 应用与练习	(12)
1.3 考虑其他人口事件影响的队列分析	(13)
1.3.1 基本原理与方法	(13)
1.3.2 应用与练习	(17)
第二章 时期分析的基本原理	(19)
2.1 时期分析的基本概念与主要目的	(19)
2.2 标准化方法	(22)
2.2.1 直接标准化	(22)
2.2.2 间接标准化	(25)
2.2.3 双标准化	(28)

2.3	假想队列方法	(35)
* 2.4	关于队列人口事件发生的提前或推迟如何 影响时期总和率的数学证明	(37)
2.5	时期频率、发生/风险率与概率	(40)
2.6	应用与练习	(43)
第三章	生命表分析	(47)
3.1	单递减生命表	(48)
* 3.2	“多递减”生命表	(51)
* 3.2.1	考虑各种“递减”因素综合影响的多递减生命表	(52)
* 3.2.2	去递减原因生命表	(56)
3.2.3	应用与练习	(61)
* 3.3	多增-减生命表	(63)
第四章	死亡分析	(73)
4.1	婴儿死亡分析	(73)
4.1.1	婴儿死亡率	(73)
4.1.2	围产期死亡率	(78)
4.1.3	新生儿死亡率与后期新生儿死亡率	(79)
4.1.4	内生婴儿死亡率与外生婴儿死亡率	(79)
4.1.5	应用与练习	(80)
4.2	死亡率的标准化分析方法	(81)
4.2.1	粗死亡率	(81)
4.2.2	按龄死亡率	(82)
4.2.3	标准化死亡率	(84)
4.3	单递减死亡生命表分析	(87)
4.3.1	如何正确估算死亡概率 ${}_nq_x$	(87)
4.3.2	关于 ${}_nq_x$ 的取值	(88)
4.3.3	很高年龄组死亡概率的估算	(91)
4.3.4	死亡率水平在个体之间的异质性	(92)

4.3.5	我国死亡率水平与人口平均期望寿命的变动趋势	(93)
4.3.6	应用与练习	(96)
4.4	多递减分死因生命表与去死因生命表	(96)
4.4.1	多递减分死因生命表	(97)
4.4.2	去死因生命表	(106)
4.4.3	应用与练习	(108)
4.5	模型生命表	(108)
4.5.1	联合国 1955 年模型生命表	(109)
4.5.2	寇尔-德曼区域模型生命表	(110)
4.5.3	里德曼-布瑞斯模型生命表	(111)
4.5.4	布拉斯罗吉特相关生命表系统	(111)
4.5.5	联合国为发展中国家编制的模型生命表	(117)
4.5.6	蒋正华等创制的中国区域模型生命表	(118)
4.5.7	如何使用模型生命表	(119)
4.5.8	应用与练习	(121)
第五章	婚姻分析	(125)
5.1	利用人口变动统计资料的婚姻分析	(125)
5.1.1	初婚量度	(125)
5.1.2	离婚、丧偶与再婚量度	(129)
5.1.3	应用与练习	(131)
5.2	利用人口普查或人口调查资料的婚姻分析	(133)
5.2.1	利用时点婚姻状态数据	(133)
5.2.2	利用回顾性调查资料	(138)
5.2.3	应用与练习	(138)
* 5.3	婚姻与婚姻解体模型	(141)
5.3.1	寇尔-麦克尼尔初婚模型	(141)
5.3.2	模型离婚表	(143)
5.3.3	应用与练习	(144)

* 5.4	婚姻生命表分析	(146)
5.4.1	初婚与离婚单递减生命表及初婚多递减生命表	(146)
5.4.2	多增-减婚姻状态生命表	(147)
5.5	婚姻挤压或婚姻市场问题	(155)
第六章	生育分析	(157)
6.1	生育量度	(157)
6.1.1	粗出生率与一般生育率	(157)
6.1.2	年龄别与年龄-孩次别生育率	(158)
6.1.3	婚后年数别生育率	(160)
* 6.1.4	生育间隔年数别生育率	(160)
* 6.1.5	总和生育率,总和婚后年数别生育率与总和间隔 年数别生育率	(162)
6.1.6	粗再生产率与净再生产率	(166)
6.1.7	生育孩次递进比	(168)
* 6.1.8	如何根据孩次递进比与孩次平均生育 年龄(或婚后年数)估算平均生育间隔	(171)
6.1.9	应用与练习	(173)
* 6.2	生育的生命表分析:初婚-生育孩次状态 生命表	(176)
* 6.2.1	初婚-生育孩次状态生命表的构造方法	(176)
* 6.2.2	应用:我国 1986 年城乡初婚与孩次别生育的 多增-减生命表分析	(178)
6.3	生育的标准化分析	(181)
6.3.1	寇尔生育指数	(181)
* 6.3.2	分离年龄构成,婚龄变动及已婚生育水平对粗出生 率影响的“三标准化方法”(双标准化方法的扩展)	(186)
* 6.3.3	“三标准化方法”的应用:我国八十年代后期 粗出生率回升影响因素分析	(189)
6.3.4	应用与练习	(194)

6.4	自然生殖能力、自然生育率与避孕效果	(195)
6.4.1	自然生殖能力的定义与不可受孕期间	(195)
6.4.2	自然生育率	(196)
6.4.3	避孕效果的量度与估计	(198)
* 6.4.4	避孕及人工流产对降低生育率的作用	(199)
* 6.5	生育模型	(203)
* 6.5.1	邦戈茨生育模型	(203)
* 6.5.2	寇尔-特拉赛尔生育模型	(210)
* 6.5.3	布拉斯的冈泊茨相关生育模型	(213)
* 6.5.4	应用与练习	(222)
第七章	迁移分析	(224)
7.1	迁移分析的时间与空间	(224)
7.2	人口迁移数据的收集	(227)
7.3	粗迁移率、流向别迁移率、流向-年龄别 迁移率以及简单迁移生命表	(229)
7.4	利用普查数据估算迁移率	(235)
* 7.5	人口迁移模型	(238)
* 7.6	应用分析:农村向城镇的迁移对我国人口 发展的影响	(241)
* 7.6.1	农村向城镇迁移人口行为转变的潜在可能性	(241)
* 7.6.2	农村-城镇迁移对我国人口增长的影响	(242)
* 7.6.3	农村-城镇迁移对我国城乡人口老化的影响	(244)
7.7	应用与练习	(246)
第八章	稳定人口模型及其扩展	(248)
8.1	经典的稳定人口模型	(248)
8.1.1	稳定人口的定义及基本原理	(248)
8.1.2	稳定人口的三个基本方程式	(250)
8.1.3	稳定人口的平均世代间隔	(254)

8.1.4	内在自然增长率	(256)
8.1.5	模型稳定人口表及其应用	(257)
8.1.6	应用与练习	(261)
8.2	静止人口-稳定人口当 $r = 0$ 时的特例	(262)
8.2.1	静止人口的特征	(262)
8.2.2	应用与练习	(264)
* 8.3	动态人口模型及其应用	(264)
* 8.3.1	描述非稳定人口动态过程的基本方程式	(265)
* 8.3.2	如何估算 $r(a, t)$	(270)
* 8.3.3	动态人口模型的应用案例之一： 估算普查间生命表	(271)
第九章	人口数据质量评估与调整	(279)
9.1	人口数据误差的类型	(279)
9.2	年龄申报误差及其评估指标与方法	(282)
9.2.1	年龄申报误差来源及类型	(282)
9.2.2	玛叶指数	(285)
9.2.3	韦伯指数	(286)
9.2.4	普查间存活比法	(287)
9.2.5	与稳定人口年龄构成比较法	(288)
9.3	如何判别人口事件数的漏报及事件 发生时间的错报	(289)
9.3.1	人口事件数的漏报	(289)
9.3.2	人口事件发生时间的错报	(292)
9.4	如何评估人口普查数据质量	(294)
9.4.1	人口分析法	(294)
9.4.2	事后核查法	(296)
9.5	质量欠佳数据的调整	(300)
9.5.1	对明显的局部误差的修正	(300)

9.5.2	图解手工修匀	(300)
9.5.3	插值法	(302)
* 9.5.4	人口模型拟合法	(304)
* 9.5.5	德曼-波特尔的人口普查年龄构成数据调整法	(305)
9.6	应用与练习	(308)
第十章	人口预测及其应用	(310)
10.1	人口预测的作用与意义	(310)
10.2	传统的单区域人口预测	(312)
10.2.1	早期利用数学函数预测人口总数增长的方法	(312)
10.2.2	分要素人口预测法	(314)
10.3	多区域人口预测方法	(320)
10.4	多区域人口预测的应用案例:我国城乡人口 动态预测分析	(324)
10.4.1	我国城-乡人口动态预测模型及基础数据	(324)
10.4.2	应用分析:逐步提高平均生育年龄对我国人口 发展的影响	(330)
10.5	应用与练习	(336)
* 第十一章	生育率与死亡率的间接估计	(337)
* 11.1	利用曾生子女存活比估算儿童死亡率	(337)
* 11.1.1	基本原理	(338)
* 11.1.2	估算儿童死亡率对应的年份	(343)
* 11.1.3	估算两个普查或调查时点之间的平均 儿童死亡率	(345)
* 11.1.4	如何估算生育率变动情况下的儿童死亡率	(347)
* 11.1.5	利用曾生子女死亡比例间接估算儿童 死亡率的实例	(350)
* 11.2	根据鳏寡比例数据估计成人死亡率 ——鳏寡法	(355)

* 11.2.1	基本原理	(356)
* 11.2.2	利用按年龄分组数据的鳏寡法	(357)
* 11.2.3	利用按婚后年数分组数据的鳏寡法	(360)
* 11.3	如何将估计的儿童死亡率及成人条件存活 概率连接成完整生命表	(363)
* 11.3.1	利用布拉斯罗吉特生命表系统的连接方法	(363)
* 11.3.2	利用模型生命表的连接方法	(366)
* 11.4	利用亲生子女法估算年龄别生育率	(367)
* 11.4.1	基本原理	(368)
* 11.4.2	应用亲生子女法估算年龄别生育率的基本步骤	(369)
* 11.5	根据两个时点平均曾生子女数估算年龄别 生育率	(373)
* 11.6	应用与练习	(380)
参考书目		(384)
专业名词术语索引		(386)
习题参考答案		(395)

绪 言

具有几千年文明历史的中国是世界上最早收集与整理人口统计数据的国家之一。然而由于种种历史的原因，直到 1982 年以前的近代中国几乎成了世界人口数据的一个空白。举世瞩目的我国 1982 年第三次人口普查以其高质量的数据揭开了所谓“中国人口之谜”。随后，全国千分之一生育调查，七省二市深入生育调查，全国百分之一人口抽样调查，全国千分之二生育节育抽样调查，全国第四次人口普查，以及其他各种各样的全国性或地方性的调查，已经并正在继续提供大量的内容十分丰富的人口数据。

然而，有了丰富的人口数据，并不等于真正认识了人口现象与人口发展过程的实质。举两个例子：（1）我国第四次人口普查主要数据公报（第四号）指出，1990 年 7 月普查时点前一年上海的死亡率为 6.36%，而宁夏回族自治区的同期死亡率却只有 5.07%。为什么社会、经济、医疗保健水平远远高于宁夏的上海市的死亡率反而高出宁夏 25.4%？（2）众所周知，美国的黑人社会、经济地位与医疗保健条件明显地比白人差。但许多年来，人口统计数字反复表明，美国黑人老年死亡率明显地低于白人的同年龄死亡率，人们称之为“倒挂”现象。同时，比此“倒挂”现象年龄较低的年龄组的黑人死亡率却都比白人高。如何解释这种美国黑人与白人年龄别死亡率曲线的交叉和高龄“倒挂”的现象？诸如此类的问题不胜枚举。《人口分析方法与应用》一书可以帮助读者理解与正确解释这些现象，找出形成各种人口现象的要素以及各个要素的贡献成分有多大。例如上海市死亡率高于宁夏的问题，可能是由于上海市老年人口占的比例较大，也有可能是由于

宁夏的死亡记录不全造成的。美国黑人老年死亡率明显地低于同龄白人，可能是由于黑人的中青年死亡率高，身体素质差的人大多数已死去，有幸进入老年期的人多数是身体素质好的人。而白人由于医疗保健条件好，不少身体素质差的人被“救”下来，活到老年期。但他们进入老年之后的死亡率毕竟要高些，致使白人老年死亡率高于黑人。也有可能是黑人老年死亡率数据不准，例如黑人有夸大年龄的倾向。人口分析方法可以帮助读者找出发生以上问题的可能的原因，并分析各种原因的主次地位。

《人口分析方法与应用》主要讲授如何运用人口学原理与数学工具来量测人口变动及分析其影响因素。它有三个主要特点：（一）综合性，即尽可能介绍我国与欧、美不同学派的人口分析方法。力图在本教材中溶欧、美、中人口分析方法与应用为一体。（二）新颖性，即力争多介绍一些近几年的最新成果。（三）应用性，即十分强调应用，每一章都有一节或数节专门阐述应用问题。要求每个学生能将所学的分析方法运用到实际人口问题的研究中去。

本课程并不包罗人口学分析方法与应用的全部。人口学理论、原理等定性分析也是十分重要的，必须认真学习，打下深厚功底，并与人口定量分析紧密结合。还有，人口学的统计分析（例如，近年来在国际人口学界发展很快的事件历史分析，包括风险模型、生存分析等）也是很重要的。同时，人口学研究还应与其他学科，如社会学，经济学，心理学，地理学等实行多学科的横向联合与交叉。这是因为人口学是一门典型的边缘学科。

本书在介绍国际、国内的人口分析的新方法与应用成果的同时，特别强调系统地全面阐述人口分析的基础知识、基本概念、原理、量测指标、模型等。该书语言通俗易懂，每节配有思考讨论题与练习，便于自学。本书既可作为人口学或与人口有关专业的研究生教材，又可作与人口有关专业的本科与专科高年级的教材。

本科、专科学生最好在修完了《人口理论教程》、《人口统计学》之后选修本课程。相关专业的本、专科生选用此教材，可选修除加注“*”号外的章节。人口学专业本科生亦可选修部分或全部加注“*”的章节。研究生则应学完包括加注“*”号在内的全部章节。

本书尽可能提供较详细的参考文献目录。除在书末列出了与本书各章或若干章有关的主要参考书目外，还在每章后面列出了与该章某些特定内容有关的补充参考文献。读者可根据课文中引用文献时给出的作者姓名及发表年份在书末所列参考书目或各章后面的补充参考文献中查出文献详细出处。

对张纯元所长、教授，袁方教授、蔡文媚教授等对本书编著工作的鼓励与支持，对涂平副教授，李涌平、郑晓瑛博士提出的宝贵意见，特别是对选修本课程的北大人口研究所八七、八八、八九、九〇、九一级研究生与青年教员蒋秉文、梁志武、李伟、王正联、刘义、蔡茜、楚军红、廖岷、赖忠文、吴德清、吴祯、董利国、李晓丽、马忠东、华民、郭柳、李忠和、于学军、王洪、康柏涛、蒋建军等表示衷心的感谢！使我终生难以忘怀的是与同学们的讨论与交流。他们敏锐的思维与质疑往往使我不得不重新思考，甚至重写原稿，为此纠正了不少错误，而且还促使我学到了不少新东西。这使我深深地体会到了教学相长的益处以及教学与自身钻研学问相互促进的奥妙。再次对同学们表示衷心地感谢！

我还对方肃平、任强、曾旭辉等同志帮助打印原稿，对蔡茜、楚军红、郭柳等同志帮助绘图与校对表示深深的谢意！

曾毅

1992年2月于燕园

第一章 队列分析的基本原理

队列分析与时期分析是人口分析两个最基本的概念。本章与下一章将分别阐述队列分析与时期分析的基本原理。

1.1 基本概念

1.1.1 队列的定义

队列的英文对应词为 cohort，也有的学者将其译为“同批人”或“同期群”。在古罗马，cohort 指三至五百人的步兵队。在生物学里，cohort 指动植物分系统的一级单位——股。在人口学里，队列 (cohort) 指在同一时期经历了同一起始事件 (event-origin) 的一群人。例如，同一年出生的人为一出生队列。同一年结婚的人为一结婚队列。又例如，同于 1977 年入校的所有同学被称为七七届校友，实际上也是一个队列。队列是以我们感兴趣的特定起始事件 (如生育，婚姻，入学等) 分类的。成为某一特定队列成员的基本条件是同一时期经历这一特定起始事件。同一时期的长度可是一年，也可是几年 (如五年)，还可以是一月甚至一天，视研究需要与数据可能而定。同一队列的成员的年龄可以相同 (如出生队列)，也可以不同 (如结婚队列，入学队列)。

1.1.2 队列的参照时间与事件的时间量度

时间观念是研究一切社会现象的最基本的要素之一。队列分

析当然也不例外。队列一般以日历年份为参照时间。例如，1952年出生的出生队列以及1975—1979年间结婚妇女的结婚队列。

队列成员经历的事件的时间量度主要有两种方式：(1) 日历年份。如1952年出生的100 000人中有299人在1955年死亡。又如1975—1979年间结婚的100 000名妇女中有350人在1980—1984年中离婚。(2) 离队列成员共同经历的起始事件的时间长度或期间(duration)。对于出生队列来说，出生后一年为一岁，出生后两年为二岁……等，这就是年龄。对于结婚队列来说，婚后一、二、三、……年为婚后期间一、二、三、……年。也有人简称婚后期间为婚龄。

年龄可分为确切年龄(exact age)与完全年龄(completed age)。例如，某女孩已出生三年零八个月，她的确切年龄为3.67岁，而她的完全年龄为3岁。完全年龄实际上是上一个生日的年龄(Age at last birthday)。确切年龄去掉小数部分即为完全年龄。同理，离队列起始事件的期间也有确切期间(exact duration)与完全期间(completed duration)之分。例如，某人已婚20年零6个月，他的确切已婚期间为20.5年，而他的完全已婚期间则为20年。

当已知队列的起始事件(如出生、结婚)发生时间及事件离起始事件的时间长度(期间)时，我们可以很容易地算出该事件发生的日历年份。反之，若已知队列起始事件发生时间与事件发生日历年份，亦可算出时间长度(期间)。例如，已知1952年出生的100 000人中有299人在3岁时死亡，那么该队列在1955(=1952+3)年的死亡人数为299人。若已知该队列1955年死亡人数为299人，亦可知该队列成员中3(=1955-1952)岁死亡人数为299人。

1.1.3 列克西斯图

人口分析中使用最早与使用最多的图示工具是列克西斯图。