

人口预测的方法与计算机应用

郭 华

一、人口预测的作用

人口的发展是人类社会经济发展中最重要的问题之一。这是因为,一方面人口是社会物质财富的消费者。在目前市场经济的体制下,今后经济发展和科教、卫生等事业的发展,都必须考虑到未来时期将要消费这些产品、享用这些设施的人口数,否则就缺少了一个重要衡量标准,无法预测未来经济市场的需求;另一方面,人口又是社会物质财富的创造者,特别是青壮年,是社会劳动力资源。预测这部份年龄段人数比例,掌握未来劳动力资源可能发生的变化,对未来经济发展具有重要意义。此外,人口又是自身进一步再生产的基础,人口再生产的周期一般在20年至30年左右,未来的人口现象和今天的人口现象存在着必然的因果关系。为了对未来的经济发展提供科学的依据,要通过人口预测了解未来的经济和人口的发展情况,了解未来人口与经济的内在联系,这就是我们之所以要进行人口预测的目的所在。同时也是社会科学的发展迫切要求方法论革命、迫切要求量化的产物。

人口预测是以人口定期统计和人口普查资料为依据,从以往人口发展的规律和现状出发,来推测未来人口的发展趋势。早在1897年,英国的经济学家马尔萨斯就着力研究经济与人口发展的关系,按照马尔萨斯的人口增长模型,推出了人口增长的指数模式。马尔萨斯认为:人口若不受食物(生活资料)的制约,将按几何数列增加,而依赖于有限土地的食物却按算术数列增加,这二者的差距,必然使生活水平下降,因而人们不得不永远贫困。马尔萨斯模型是建立在人口可以无限增长这个假说上。因此他的预测结果具有很大的片面性。

1844年,比利时数学家P·F·Verhulst创建了逻辑斯蒂曲线,1923年美国P·Pearl与L·J·Reed把它用于人口研究。根据这个方程提出的假设,可以得到人口增长的逻辑斯蒂方程,这个方程克服了马

尔萨斯提出人口增长模式中人口数会无限制地增长的弱点。但由于逻辑斯蒂方程有一个最大人口限制,用它来预测人口,结果都偏低。

预测人口发展趋势有各种不同的方法,每一种预测方法都有其自身的特点,我们要进行人口预测,必须根据不同情况建立正确地人口模型。掌握正确地人口预测方法,是人口预测的关键。

二、人口预测的方法

人口发展过程的定量预测,是根据现有的人口统计数据 and 历史资料,从当前实际的人口状况出发,按照合理的控制要求和假定,应用科学的方法,对未来人口发展趋势进行几年、几十年甚至上百年的预测。预测内容,除了预测计划期各年度的人口总数外,还要预测人口的性别、年龄、地区构成,预测人口的出生、死亡和自然增长率的变化以及未来的人口构成中劳动力状况、抚养水平、老化水平、平均预期寿命、平均年龄、年龄中位数、人口密度等各项人口指标。

人口定量预测的方案很多,但基本方法可分为数学法和要素法。①数学法主要是利用数学模型对人口总数进行预测。直接通过一次和多次普查的人口数、增长率及增长量来推测未来人口发展。它包括人口发展方程、逻辑斯蒂曲线方程,线性方程和多项式方程等等。这种方法概念简单,使用方便。但只能机械、粗略地对总人口进行预测。所以目前更多地是使用要素法。②要素法主要是通过影响人口变动的三个要素——出生、死亡、迁移分别进行预测,从而对分年龄、分性别人口数进行预测。

在用要素法进行人口预测时,又分为静态预测法和动态预测法。静态预测法的计算公式和假说的初始条件都比较简单,人口发展的参数(如死亡率、出生率等)也是固定不变的。因此,预测中很难把未来人口发展随时间变化的规律考虑进去。例如,年龄移算法的公式如下:

$$\text{预测期总人口} = \text{预测期初人口数} + \text{预测期内出生人数} - \text{预测期内死亡人数}$$

预测时,用固定的年龄别生育率、年龄别死亡率对预测期内出生人数和死亡人数进行列表推算。

若要用1990年的人口数,预测2000年的人口数,则年龄移算法的公式为:

$$\begin{matrix} 2000年 \\ 的人口数 \end{matrix} = \begin{matrix} 1990年 \\ 的人口数 \end{matrix} + \begin{matrix} 1991年到 \\ 2000年出 \\ 生的人数 \end{matrix} - \begin{matrix} 1991年到 \\ 2000年死 \\ 亡的人数 \end{matrix}$$

它不考虑怎样出生、怎样死亡的过程,因此说它是静态的。这样的人口预测,随着时间的推移,积累误差会越来越大,影响人口预测的精确度。所以这种静态预测方法,只能用在预测年限比较短,预测精度要求不高的情况下。

另一种是动态预测法。动态法首先是要根据生育、死亡等建立动态的数学模型。与静态法的区别在于,将出生、死亡、迁移等都看成是随时间变化而变化的量。一个较为常用的动态数学模型是确定型模型。当研究的地区人口数量很大时,可以把人口的出生、死亡、迁移的变化看成是一个连续变化的量,从而从离散模型导出连续的人口发展方程。

设 ΔX 是一个微小的年龄区间, Δt 是一个微小的时间区间,年龄从 X 变到 $X + \Delta X$,时间从 t 变到 $t + \Delta t$,当 $\Delta X, \Delta t$ 趋于零时,不难导出连续的人口增长数学模型:

$$\begin{cases} \frac{\partial n(x,t)}{\partial x} + \frac{\partial n(x,t)}{\partial t} = -q(x,t)n(x,t) \\ \quad + W(x,t) \\ n(x,0) = n_0(x) \\ n(0,t) = \phi(t) = F(t)N(t) \end{cases}$$

其中 $q(x,t)$ 为 Δ 时刻 t 、年龄为 x 的死亡率
 $W(x,t)$ 为 Δ 时刻 t 、年龄为 x 的迁移人数
 $\phi(t)$ 为时刻 t 的出生人数
 $N(t)$ 为时刻 t 的总人口数,

这是一个一阶线性偏微分方程,在人口预测的实际计算中,不是求它的通解,而是化离散模型进行迭代计算。

确定型模型有两个优点,①生育率、死亡率、迁移率等参数可以随时间变化,能够看出人口变化的动态情况。②按年龄的组成可以灵活处理(如可以按婴儿、幼儿、学龄、劳力、老年等分组),易于进行人口分析。但它要求基础数据正确,不能有

大的误差,看不出基础数据误差对后来的影响。

再就是随机模型,这个模型就是注意到人口发展变化是一个非常复杂的过程,因为具体考察一个人在当前死与不死、一个育龄妇女当前生育与不生育来说,都是不确定的,即有一定的随机性。人口的出生率参数和死亡率参数每时每刻都在变化。因此反映人口发展过程的各因素都是随机变量。这种预测的结果可能更符合客观实际。

三、人口预测的软件

随着人口学研究水平的提高,在人口状况分析中,越来越广泛地使用定量分析,人口统计学理论也不断发展完善。但是人口统计学的进步同时带来了一些新的问题。现代人口统计理论和统计技术日趋精细复杂深奥,使得一般实际工作者难以用大量时间去专门学习掌握这些理论和方法;同时,也难以通过长时间、大规模的繁琐运算来取得某一项具体预测结果,更谈不上耗费大量人力物力去编制多种用于比较研究的预测方案。这就极大地妨碍了这些科学理论技术的推广应用。在这种情况下,运用现代计算机技术把现代社会科学定量研究的最新理论成果变成实用型科学管理和科学预测工具,以利于各方面的应用,就成为一些十分重要而有意义的工作。

在70年代末、80年代初联合国人口基金会就根据世界一些有代表性和典型的国家生育率水平,建立了一些比较常用的数学模型,即西模型、东模型等,编制成各种计算机软件,供研究人员应用。这套软件是用FORTRAN语言编制的,研究人员根据各地的实际情况,选用不同的模型,输入各类数据和参数,通过计算机处理,可以计算出各年的人口年龄、性别结构,并计算出未来各年的平均年龄、年龄中位数、模型生命表及画出人口金字塔等人口数据的图表。

80年代初,我国的控制论专家宋健等人将控制论的理论应用到人口预测领域,也用FORTRAN语言编制了人口预测模型,用这个模型预测了我国100年人口发展情况,并提出了我国人口发展的战略部署。

到80年代中期,联合国人口基金会和国际劳工组织又在原有的基础上对一些人口软件进行不断完善和修改,并推广应用这些统计预测软件。这些软件的生育模型采用了胎次法来预测,使得模型更加合理,且统计模型方面也增加了很多根据发展中国家的实际情况而制定的模型。这些软件多采用BAS

IC语言编程,给操作者带来了更多的方便:①加强了编辑功能。因为FORTH语言对数据输入的格式要求很严格,在录入数据时严格按数据格式录入很麻烦,不利于一些特殊数据的录入(数值太大或太小)。②提高了人机对话功能。因为用FORTH语言的模块调用功能很严格,不利于发现问题及时修改数据,人机对话功能不强,特别是对这些软件不熟悉的人,就无法操作。③增加了预测内容。目前的这些软件,可以根据研究对象不同,随意预测人口问题的某几个方面的内容,并增加了人口迁移的因素。在图形功能方面,除了可显示人口金字塔外,还可用直方图、曲线图来描述。这样使得人口预测软件更加完善,更加便于操作。

这些人口预测软件应用,给研究人员的研究工作带来了很大的方便。但因为预测模型的规范化,只能按典型模型去模拟某一地区的人口发展状况时,某些条件不满足时,也可能会出现一些偏差,所以这些预测软件一般较实用于预测范围比较广、时间为中长期的情况。

四、自编人口预测软件应用

由于模型软件存在着以上的不足,为了使预测结果更能符合实际,许多地区和部门也采用自编人口预测软件。武汉大学人口研究所也根据研究需要,用BASIC语言和DBASE数据库,编写了人口预测软件,并在“湖北省人口及人力资源发展战略”的项目和“秭归贫困县人口发展调查”项目中得到应用。

在这些自编程序中,我们采用的是,

①树型结构的分级菜单管理程序。在程序中,我们采用3级管理。首先是主菜单,包括预测、修改、打印和退出这四个功能,而4种功能模块,又分别调用各自操作的子模块。这些操作子模块又控制程序的基本数据、结果数据和程序之间的调用,从而成为一个完整的系统。②数据库的应用。在程序中,我们除了建立了基年资料的录入数据库外,还有一些中间运行过程中的数据库和最后结果的数据库,这些数据库的应用和有机地结合,使我们地操作更加方便。③数据模型的合理建立。在程序中,我们根据某省或某县的实际情况,选择了动态的生育模型,并按各自生育模式曲线、用正弦函数、阶梯函数、指数函数等初等函数和正态分布、F—分布、T—分布、 X^2 —分布等概率分布对各类地区育龄妇女生育模式进行拟合、分析,找出较有代表性的生育模式,提高预测结果的准确性。④输出结果直观。在程序中,我们尽量考虑能很直观地看出人口发展的情况,所以输出形式为图表合一,各种人口变化趋势的图表从输出屏幕上和打印机上输出,一目了然,增加了输出结果的可读性。

总之,我们用这个程序并根据某省、某县的实际进行人口预测,为湖北省政府提供《湖北省人口发展与人力资源的发展战略》的研究报告和为贫困县总课题组提供的《秭归县的人口发展状况》的研究报告,引起了政府部门和学术界的高度重视。这也是利用计算机进行社会科学研究一个典型实例。

(责任编辑 徐云鹏)

(上接第30页)综上所述,农民负担是沉重的,后果是严重的。农民负担过重的原因是多方面的,有浅层次的,也有深层次的;有经济的,也有政治的、法律的;有政府的原因,也有农民自身的原因;有上级部门的原因,也有基层的问题;有人的因素,也有体制的制约。应该仔细分析,抓住主要矛盾,下狠心从根本上解决农民负担问题。

注释:

- ①许宝健:《何以减而不轻——农民负担问题透视》,《经济日报》1993年2月26日。
- ②④庹震:《为了农民的合法权益……》,《经济日报》1992年12月16日。
- ③黄发源:《方方面面说负担》,《湖北日报》1993年2月21日。
- ⑤⑥许宝健:《怎一个“乱”字了得——农民负担问题透视》,《经济日报》1993年2月19日。
- ⑦许宝健:《高度重视农民负担过重问题》,《经济日报》1992年12月29日。
- ⑧许宝健:《警惕农业重新陷入徘徊——人大代表朱同炳坦陈农业深层矛盾》,《经济日报》1992年4月4日。
- ⑨见《经济日报》1993年5月13日。

(责任编辑 王雪松)