

农户在牧区是牧户)。这许多家庭可说是这个经济区域的细胞。细胞之间通过经济分工合作,交流和服务而聚合成不同的社区集体,在行政系统里即村和嘎查等。若干社区集体中涌现出常说的第三产业如贸易、金融、餐饮、修理等服务性行业,即成为经济较发达的各个层次的中心,一般称作市镇。一个经济中心拥有它所服务的许多细胞,这些就是它的腹地。经济区域是有不同层次的,越发达的经济中心,它的层次也越是往上升级,它包括的地域也跟着扩大,拥有的腹地也越大,包括的基本经济细胞也越多。

从这些构成经济区域的条件来看具体的赤峰市,要进一步发展,交通设施是首先要加以推进的重要条件。通过比较先进的交通设施才能把已经在自然经济里形成的经济细胞,即农户和牧户,联成具有适当服务体系的社区集体,更把它联系成更高层次的市镇,在这些市镇中发展工商业和第三产业,更形成一个地区更高一级的经济中心,相当于现在的拥有几十万人口的中等城市。当经济发展再进一步时,可以预期在自治区的范围内出现一个拥有百万人口的大城市,成为自治区的经济中心。

赤峰市今后在这个经济区域的格局中取得什么地位,要看今后能达到什么样的经济水平。以我个人初步猜测,在自治区的西部地区可能形成一个一定层次上的经济区域。至于这个经济区域的中心在那里,现在尚难预测,但是现在的赤峰市,如果上述的铁道系统能够实现,可能是这个经济区域中心的优先选择。

还有一个问题应当及时考虑的是内蒙古自治区作为一个经济区域来看,还需要找到一个为它服务的出口海岸,就是一个向外吞吐的港口。从赤峰市打算,能越靠近这个港口越有

利。究竟怎样选择则必需从国内的总格局来研究。在目前我只能提出这些问题,以供研究和及早考虑,还不能作出具体的建议。

我是一直关心边区,特别是少数民族地区经济发展的。这次来赤峰访问,一方面得到了很多新的知识,知道了以前不知道的事实,更重要的,受到深刻教育的,就是发现了许多我还无法作答的问题,推动我进一步研究和思考。我个人不容否认的是已经接近衰老了,能继续从事调查研究的机会越来越少了。但是我感到高兴的是已有年轻的接班人可以继续在我提出的问题上更深入地进行研究下去。我们一定会更加努力地工作,为边区的经济发达作出进一步的贡献。

(作者单位:北京大学社会学人类学研究所 1995年7月25日)

☆☆☆☆☆☆☆☆☆☆

关于建立中国民族社会学 数据库的设想和倡议

郑 凡

一、设想的由来

欣闻中国民族社会学研究会在京成立!我们对自己正在进行的一个云南省重点课题《民族社会学概论》,也更增添了研究信心。我们在课题研究的过程中深切感到,一方面,现时代的民族问题与和平、发展主题交织在一起,



意义重大；另一方面，真正独立的民族社会学，目前在世界大多数国家、在现代学术史上似乎都处于长期停滞不前的境地，尽管一些邻近的学科如社会学、文化人类学、民族学、国际关系学的相关研究，为它积累了大量的经验材料和某些理论线索，但是中国民族社会学的完整框架，还有待于系统性的创建。

任何一种成系统的学科建设，总是需要众多的同行来参与，需要开展广泛的学术交流，因而，我们对新近成立的研究会寄以厚望，籍此向全国的同行们提出一项倡议：如同后发展国家的现代化可以在某些技术领域走捷径，中国的民族社会学或许也能在研究手段“后发制人”，赶上即将来临的 21 世纪信息革命新浪潮，借助电子计算机的、乃至多媒体信息网络的应用技术，建立一个大型数据库 CERS (Chinese Ethnosociological Research System, 即“中国民族社会学研究系统”)！

我们注意到，在中国社会科学界，正是某些后起的研究领域却率先做了类似的尝试，如 95 年 6 月间见诸专业刊物和内部交流材料的《中国妇女研究文献数据库》、《世界妇女研究文献数据库》，它们由全国妇联研究所建立，已录有数以万计的文献检索条目，内容广及政治、经济、法律、文化、教育、就业、家庭、环境、健康、人口、妇女学科研究等，此外还有一种现象值得注意，由市场经济的特点所决定，国内的工商界已在一些方面走在研究界前面，笔者见到一种光盘数据形式的（中国百科全书）（SunSoft 公司出品），其中接近于民

族社会学需要的，有专门介绍中国 56 个民族一般历史和习俗的图文信息，不过在笔者看来，也正因为商业特点，有关的民族资料几乎是“趣味性”的，与我们的专业研究需要还相距甚远。所以，我们所设想的 CERS,还只能靠全国的民族社会学研究者们自己来创建。

参考各种已有的实例，可以肯定，民族社会学数据库及其联网运用，①能够以系统性和开放性的优点，为研究者储备巨大的信息资料，对国内外各民族的经济、文化、人口、政策问题和调查统计数据等，加以大范围的跟踪利用与处理；②能够利用一般数据库软件所提供的简单计算和宏命令公式计算，或链接起有关的统计软件，进行定量研究所需要的大型统计分析；③能够使研究课题在学术界获得迅速有效的协作与交流，使研究成果在社会公众中得到加速传播。

二、试以民族人口模型做示范

考虑到我国社会科学界利用信息技术的一般水平比较低，许多部门迄今只把计算机当作“打字机”来用，让我们虚拟一个演示实例——民族人口的预测模型。设想一位研究者坐到计算机前，按照简单的操作指令，从数据库 CERS 中，查询 YEP (Yunnan Ethnos' Population, “云南各民族人口”)分库的记录——



中国·云南·民族人口规模的变化和比较

(ERS — YEP/1953~1987 年部分记录)

单位：万人

年 份	1953 年	1964 年	1978 年	1979 年	1980 年	1981 年
少数民族	563.49	640.25	948.42	963.88	983.15	1002.22
汉 族	1149.79	1410.70	2143.05	2170.91	2190.24	2220.55
全省总数	1713.27	2050.95	3091.47	3134.79	3173.39	3222.77
年 份	1982 年	1983 年	1984 年	1985 年	1986 年	1987 年
少数民族	1032.17	1046.13	1065.16	1085.15	1110.96	1135.39
汉 族	2250.89	2272.92	2297.11	2321.07	2244.68	2377.61
全省总数	3283.05	3319.5	3362.27	3406.21	3455.64	3513.00

注：1953、1964 年数为普查年中数，其余年为人口年报年末数，故本表 1982 年数超出该年度人口普查数。资料来源：《云南四十年》（云南省统计局编，中国统计出版社，1989 年 7 月）。

当然，数据库并不限于被动显示已有的记录材料，在电子化信息的意义上，它既是一种资料库，同时还提供了一种内在于资料结构中的信息加工场，能够辅助人们作进一步的研究，例如就依据上表的人口数据作定量处理，我们来预测云南省少数民族和汉族人口的发展，这在理论研究和对策研究上都是有益的。

熟悉人口统计学的人知道，人口预测分为线性模式、马尔萨斯模式、逻辑斯蒂方程和粗率法、年龄移算法等等，它们各有一定的适用特点。这里让我们用两种方法来做些尝试。上表的云南汉族和少数民族人口数据，其起始点是 1953 年，然后不完全连续地一直记录到 1987 年，其中包含了我国第一、二、三次人口普查的年份，受资料来源本身所限，1990 年第四次人口普查的有关数据未列入其中，但在这里正好可以用 1987 年以前的数据来验证，看更适合描述云南汉族和少数民族人口发展趋势的数学方法，是线性模式还是马尔萨斯模式，然后再选用较好的模式去预测未来。

人口预测数学方法的线性模式，其基本假

定是人口增长速度与人口的原始规模成正比，并且各个阶段的增长速度之间没有大幅度的波动，因而以时间为自变量、人口数为变量，两者间的相关关系大致呈现直线性的因果联系，于是有如下的方程：

$$\frac{d p}{d t} = r p_{07}$$

该方程解的形式为：

$$p(t)=a+bt$$

式中 $p(t)$ 表示 t 时刻的人口规模。根据不同时期的人口规模数据，可用最小二乘法做曲线拟合求出 b 和 a ，然后外推任何时刻的人口数。

我们用这样的方法分别建立云南汉族和少数民族人口发展的线性模型，代入上表 1953~1987 年的人口数据，进而分别“预测”它们到 1990 年时候的人口数，再用 1990 年第四次人口普查的实际数据作检验，结果显示：云南汉族人口发展的线性模型与实际情况相比偏差较小，最后计算结果偏差约为 25%；而



同样的线性模型用于云南少数民族人口数据的计算，最后偏差高出汉族一倍多，达到53%，这是为什么呢？且参考有关方面的定性说明，“建国以来，云南少数民族人口的发展，在总趋势上与汉族和世界上各民族的趋势是一致的，基本遵循低——高——低的过程，只是各种政策对其人口总量的变化影响更大。更改民族成分，边境一线少数民族有规模地出境、回国的事时有发生；特别是在严格执行计划生育政策后，汉族人口发展已经明显地带上了控制的痕迹，而少数民族无论从人口增长势头，还是生育模式上，人为控制的作用相对不大。由于这些原因，云南少数民族人口总量的波动幅度更大，发展相对迅速。”^[1]由此我们能够肯定两点：①定量模型的研究也反映出，1953年以来云南少数民族人口的发展，由于历史条件和政策效应等方面的原因，比之云南汉族人口的发展有较大幅度的波动；②与此相应，云南少数民族人口发展的预测可能需要用非线性的模型来加以处理，例如借用上文提到的马尔萨斯模式。

马尔萨斯模式的基本假定也认为，人口增长速度一般与人口的规模成正比，不过由于可能存在的其它相关变量的干扰，时间变量与人口变量之间可能呈现非线性的、即某种曲线型的因果联系。表现在人口公式上，方程的解的形式是：

$$p(t)=ce^{rt}$$

计算时，对不同时刻的人口规模序列作对数变换，仍以最小二乘法求出各项人口系数，并外推任何时刻的人口增长数，以此方法，我们来建立云南少数民族人口发展的非线性模型，代入1953~1987的人口数据，最后证明这样做的误差只在15%左右。按照数理统计的检

验准则，云南少数民族和汉族人口发展模型的有关计算，其相关系数值分别达到0.993和0.985，两者都大于显著性标准，证明该模型有一定的可信度。下一步就能得到我们这项定量研究的真正预测结果了——以云南少数民族人口发展的非线性模型和云南汉族人口的发展线性模型，分别预测它们到本世纪末即2000年的人口数，以及有关的人口比例，大致结果如下：

- 1.本世纪末，云南少数民族的总人口数可能超过1,500万；
- 2.届时，云南汉族的总人口数可能达到2,900万上下；
- 3.由此，云南少数民族与汉族人口之比，可能达到52:100左右；
- 4.云南少数民族占全省总人口的比重将可能继续上升到35%上下。

我们的这些预测，也能够间接地得到其他一些研究结果的支持。如云南的不少人口专家都曾定量或定性地预言，云南的少数民族人口比例将继续增高，而对云南省本世纪末的总人口数，有的预测为4,300；有的根据妇女的总和生育率并附加了体现政策效应的控制变量，建立另外的预测模型，得出4,456万、4,337万和4,280万三种可能方案下的人口总数^[2]。这些都接近于我们以上所述，即云南2000年时人口总数达到4,400万左右，以及少数民族人口的比例将继续增高。

假如有人这样问：上述的定量研究也可以在计算机数据之外进行，你们这个CERS又特殊在哪里呢？就只是通常所听到的，计算速度比较快吧？让我们回答说，是的，数据库依托于计算机和高速运转，尤其在显示和处理巨量的数据信息时体现出它的极大优势，然而又不



仅如此，数据库资料和专用程序之便于公共利用，同样是它值得强调的重要特征。比如说，考虑到“云南各民族人口”(YEP)数据库的利用者不一定都具备人口统计学的专门知识，或没有必要一次次地重复思考和演算上述的人口发展模型，我们可以将上述的(以及其它的)模型演算过程转化为储存在数据中的固定程序(利用数据库宏命令演算公式，或一定变量域与Excel/Lotus等软件、SPSS统计软件等的程序链结)，今后就让计算机自动地向查询者提问：

您需要按哪一种人口发展模型来预测云南汉族/少数民族人口的未来发展？

——请选择 A.线性模型(适用于人口发展波动幅度较小的情况)；B 非线性马尔萨斯模型(适用于人口发展波动幅度的较大的情况)；C.……；D.……

您需要预测未来几年的云南汉少数民族人口总数？

——请选择：X.未来3年；W.未来5年；Z.未来10年；……

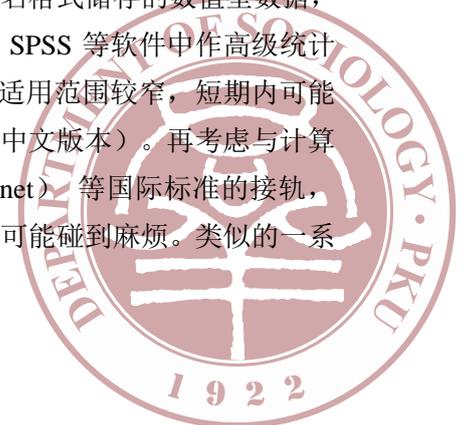
这样一来，只要数据库的使用者输入简单不过的字母代号，数据库就将按照所给定的参数条件和预先编好的运算程序去工作，很快把预测的结果显示出来。

联系到我们这个CERS数据库系统的建立和使用，民族社会学的理论研究也会显得更有意义。在初创阶段，究竟为数据库设置什么样的人口发展模型(运算程序)？对云南各民族未来人口的预测究竟用哪一种方式更为可靠？我们的上述研究只不过是一种预演。在数据库的使用和改进过程中还需要特别注意，各民族人口发展的数据模型必须随着实际的变化而变化，即使证明了云南少数民族的人

口发展从1953年至少比较接近于非线性的规律，它却不一定永远如此，而很可能在有关因素变化了的情况下，波动幅度趋于缓和，于是变得接近于云南汉族的情况，可以改用线性发展的模式来拟合其轨迹；再者，从统计学和预测科学的观点看，由于预测总要面对一定程度的不确定性，无论对云南少数民族还是汉族的人口发展作预测，我们往往需要同时选择几种不同的人口模式来演算，使之有所比较和参照。

三、让数据库系统促进学科共同体的成熟

数据库的运用，还涉及文字、数值、声音、图形、图象、动画和视像资料的多种处理方式，现有的数据库软件包又分单用户的、网络用户的、汉化和未汉化的许多种。不过我们这里最关心的是，不断出新的多媒体信息技术，一方面为我们提供了从事民族社会学数据库建设的丰富手段，另一方面则把解决现代信息技术与传统“治学”方式的矛盾提到了日程上。仅举其一点来谈，民族社会学领域各自为政的人员、机构和研究方式，就需要想出办法来适应数据系统的标准要求，后者是高速率、低冗余的现代化信息传输所必备的条件。例如，数据库字段名(即索引项)的统一值得考究，若使用中文版数据库软件，好处是字段名能够上汉字，然而以中文段名格式储存的数值型数据，却不能再直接调入SPSS等软件中作高级统计处理(因为SPSS适用范围较窄，短期内可能看到双字节内核的中文版本)。再考虑与计算机互联网络(Internet)等国际标准的接轨，中文字段名格式也可能碰到麻烦。类似的一系



列技术标准化要求，提示我们要长远地考虑中国民族社会学数据库的建设，从组织条件上讲，或许需要由少数有实力的单位来牵头，或依靠全国的学会对各地作出统一的协调布置。

这一来，民族社会学数据库系统的建立，实际上从一个侧面对民族社会学学科共同体的成熟提出了紧迫和要求。我们看到的现状是，民族社会学的理论框架在中国还处于初创，或依据国内外的一些线索进行重组、再阐释或再创造的阶段。相应的，与其说目前的主要难点是理论还不系统，应用的范围还不确定，不如说研究者之间的交流更成问题，被动地、传统方式地互相阅读已公开发表的论文，还是主要的甚至唯一的交流手段，包括本文作者在内的一些研究者，还处于“打一枪换一个地方”的散兵游勇状态。

笔者曾经在一篇文章里提及对中国社会学发展趋势的一种预见，认为：“本世纪初和70年代末两度引进西方理论以来，中国大陆社会学的历程极其曲折，但是从现在的一些迹象看，它似显示出研究范式(paradigm)和范例(examples)趋于成熟的走向，其标志是，把引进之后脱离西方原有背景的概念术语用到具体课题中，在用法上经历一种逐渐形成新背景的过程，在其中，‘西’与‘中’、经典研究和新的研究有可能相互参证，积累起中国课题的研究例证（有的即成为范例），并以一定的理想研究准则逐渐获得中国社会学家们的认同（即形成范式）。”^[3]我们猜想，这将是中

国的社会学——也包括民族社会学在内，更有意义也更具难度的一个发展阶段，因为“科学尽管是由个人进行的，科学知识本质上却是集团产物，如不考虑创造这种知识的集团特殊性，那就既无法理解科学知识的特有效能，也无法理解它的发展方式。”^[4]一个学科共同体的研究范式，是研究者们对于本学科基本概念和方法的共识，好比统计学中的“概率”和“显著性判别标准”那些基础概念及其演算程序，民族社会学数据库系统中需要加以统一的框架设计、索引条目、检索方式和处理程序，它们的相对的结构化和标准化，大约也标志着民族社会学学科共同体的成熟。

注 释

[1] 见《云南四十年》云南省统计局编，中国统计出版社，1987年7月。

[2] 见云南省第四次人口普查科学讨论会论文：张传富、张建华“云南人口现状和发展问题综述”；常建军“云南人口发展近期态势分析”；丁玉琨“未来云南人口发展趋势的预测”等。

[3] 郑凡：“中国社会学、人类学在生育健康中的应用”，载《云南社会科学》，1994年第4期，45~48页。

[4] T.S.库恩：《必要的张力》，纪树立等译，福建人民出版社，1981年，第XII页。

（作者单位：云南省社会科学院社会学研究所）

