

科学的方法应予以科学的应用^{*}

——与童星、刘松涛商榷

尹海洁

中国从1979年恢复重建社会学以来,以社会统计学为主要内容的现代社会学研究方法的应用受到了普遍的重视。在这方面中国社会学学者进行了不断的尝试,统计分析方法的运用从简单到复杂,使得中国社会学在研究方法的应用上与西方发达国家的差距越来越小。但中国的社会学毕竟起步较晚,因此在现代统计分析方法的应用上也还存在一些不尽人意之处。本文仅就童星、刘松涛二位学者发表于《社会学研究》2000年第4期上的《城市居民最低生活保障线的测定》一文(下称童文)中在方法的应用上存在的一些问题,与童星、刘松涛二位学者进行商榷。

一、正确的研究结果不应该自相矛盾

一项科学研究工作的结论性观点,应当获得相应事实的证实或者得到科学的实验或检测的支持,这是我们进行科学研究工作所应遵循的基本前提。如果所提出的观点与检测或计算的结果不吻合,甚至相矛盾,则只能是观点不正确或者是检测或计算方法使用不当。童文恰恰在这一基本原则出现了问题。

童文中用了101个城市最低生活保障线数据与国内生产总值等5项总量指标和人均国内生产总值等5项平均指标分别进行了一元线性回归分析。结果在5个总量指标回归模型的決定系数调整值(Adjusted R Square)中,最大的仅达到0.369。5个平均指标回归模型中,除平均工资回归模型的拟合度稍好以外(Adjusted R Square为0.566),其他4个模型的決定系数调整值均低于0.323,拟合情况均不够理想(见童文表1和表2)。为此,作者在原有单一指标模型的基础上加进两个社会性指标,即城市级别和地区分布,并用总量指标的国内生产总值和平均指标的人均国内生产总值分别进行了综合指标的多元线性回归分析。以国内生产总值为例,童文对加进两个社会性指标以后的模型的评价为:“判断该模型拟合优度的判定系数R Square的值为0.666,调整值为0.655,说明线性度很好,或者说,当用国内生产总值、地区分布和城市级别三者共同预测最低生活保障线标准时,三者能够‘解释’最低生活保障线变异性的66.6%。它比用国内生产总值一个指标预测最低生活保障线时的決定系数0.376增加了0.290,说明多元回归模型的解释力度大大提高。而且,多元回归模型的标准差为14.8931,比简单回归模型的标准差20.2276低了5.3345,说明多元回归模型的预测精确度提高了”(童星、刘松涛,2000:48)。对于以人均国内生产总值、城市级别和地区分布为自变量的多元回归模型,童文的评价基本与上相同。

为证明模型的有效性,作者使用江苏省12个城市的15个最低生活保障线数据对文中给

* 本文在写作过程中曾得到中国社会科学院社会学研究所许欣欣博士的指点。

出的所有回归模型进行了检测。尽管作者认为：“本文所述的回归模型(除了平均消费百分比法之外)在对江苏省 12 个城市的检测中被证明是适用有效的。”但是作者却没有注意到，如果把用各个指标的一元回归模型计算的预测结果与用其对应的多元回归模型计算的预测结果相比较的话，预测效果的优劣恰好与前面给出的模型的优劣完全相反。预测结果如表 1 所示。表 1 中的预测值 1 是用国内生产总值一元回归模型计算的结果；预测值 14 是用国内生产总值的多元回归模型计算的结果；预测值 6 用的是人均国内生产总值一元回归模型计算的结果；预测值 13 是用人均国内生产总值的多元回归模型计算的结果。从表 1 中可以看出，对总量指标的国内生产总值来说，用一元回归模型计算的预测值 1 比用多元回归模型计算的预测值 14 更接近于观测值(低保值)；对于平均指标的人均国内生产总值来说，用一元回归模型计算的预测值 6 也比用多元回归模型计算的预测值 13 更接近于观测值。

表 1 回归模型的检测表(1)

城市(年份)	低保值	预测值 1	预测值 14	预测值 6	预测值 13
南京 1996	120	126.29	145.16	123.79	147.66
南京 1997	140	129.68	147.29	126.71	149.48
苏州 1996	140	137.56	134.47	131.47	139.15
苏州 1997	180	140.98	136.31	134.22	141.05
无锡 1996	140	132.66	135.85	134.68	136.41
无锡 1997	155	136.42	138.20	138.69	138.91
常州 1997	140	121.27	128.73	126.60	131.37
镇江 1997	120	117.96	126.65	126.41	131.25
扬州 1997	120	118.51	127.00	118.65	126.41
南通 1997	120	124.70	130.87	116.79	125.26
泰州 1997	120	116.53	125.76	115.43	124.41
淮阴 1997	100	112.37	132.16	111.55	121.99
盐城 1997	100	119.55	127.65	113.69	123.32
徐州 1997	100	123.65	130.21	115.27	124.31
连云港 1997	100	113.65	123.96	113.70	123.33

说明：低保值是最低生活保障线；该表的低保值、预测值 1 两栏取自于童文表 10 中的相应栏目；预测值 6 取自于童文中表 11 的相应栏目；预测值 14 和 13 两栏取自于童文表 13 中的相应栏目。

如果用预测值与观察值的相差值来比较，这个特征则更明显，相差值见表 2。表 2 中相差值 1 是用国内生产总值的一元回归模型计算的预测值 1 与观察值(低保值)的差；相差值 14 是用国内生产总值的多元回归模型计算的预测值 14 与观察值的差；相差值 6 是用人均国内生产总值的一元回归模型计算的预测值 6 与观察值的差；相差值 13 是用人均国内生产总值的多元回归模型计算的预测值 13 与观察值的差。对表 2 中的 4 个相差值分别计算平方和可得： $\sum[\text{相差值 } 1]^2 = 249.26$ ； $\sum[\text{相差值 } 14]^2 = 405.62$ ； $\sum[\text{相差值 } 6]^2 = 242.16$ ； $\sum[\text{相差值 } 13]^2 = 339.7625$ 。这几个数值表明，不论是总量指标还是平均指标，用一元回归模型计算的预测值与观察值之间的离散程度都远远小于用多元回归模型计算的预测值与观察值之间的离散程度。这说明，童文中的一元回归模型的预测效果远远优于其加入两个社会性指标后的多元回归模型。这个结果恰好与前述的多元回归模型的预测精度比一元回归模型的预测精度高的结论相反。

表 2

回归模型的检测表(2)

城市(年份)	相差值 1	相差值 14	相差值 6	相差值 13
南京 1996	6.29	25.16	3.79	27.66
南京 1997	-10.32	7.29	-13.29	9.48
苏州 1996	-2.44	-5.53	-8.53	-0.85
苏州 1997	-39.02	-43.69	-45.78	-38.95
无锡 1996	-7.34	-4.15	-5.32	-3.59
无锡 1997	-18.58	-16.8	-16.31	-16.09
常州 1997	-18.73	-11.27	-13.40	-8.63
镇江 1997	-2.04	6.65	6.41	11.25
扬州 1997	-1.49	7.00	-1.35	6.41
南通 1997	4.07	10.87	-3.21	5.26
泰州 1997	-3.47	5.76	-4.57	4.41
淮阴 1997	12.37	23.16	11.55	21.99
盐城 1997	19.55	27.65	13.69	23.32
徐州 1997	23.65	30.21	15.27	24.31
连云港 1997	13.65	23.96	13.07	23.33

说明:相差值 1 取自于董文表 10 中预测值 1 中的相差值;相差值 14 取自于董文表 13 中相差值 14 一栏;相差值 6 取自于董文表 11 中预测值 6 中的相差值;相差值 13 取自于董文表 13 中的相差值 13 一栏。

二、假设检验的应用只有在随机抽样的前提下才有价值

数理统计的发展为人们从更大的范围上认识事物提供了方便、快捷、科学、准确的方法。推论统计是数理统计的核心内容,它解决的问题是如何通过对少量样本的了解来达到认识总体的目的。假设检验是推论统计的重要内容之一,它的目的是通过构造适当的统计量,利用样本统计量或样本状态对总体参数或状态作出合理的推断。但使用假设检验有一个非常重要的前提条件是,样本必须是通过随机抽样获得的。也就是说,只有对随机样本,统计推论、假设检验才有价值。另外,假设检验并非是检验样本,而是针对样本情况检验总体的对应项是否显著。对于非随机抽样的样本,其调查结果是不可以用统计推论的方法推论总体的。因此,对其进行任何假设检验不但毫无意义,而且容易造成无推广价值的结果被误用。关于这一点,很多统计学书中都有明确的论述。卢淑华在她的《社会统计学》中是这样阐述的:“对于统计推论,还有一个基本的,同时又是不言而喻的假定,那就是抽样必须是随机抽样。这一点在自然科学中实现起来并不难,而在社会调查中或由于总体清单找不全,或被访人难以找到,从而变通成碰到谁就调查谁的偶遇以及由熟人再找熟人的滚雪球式的调查。应该指出,这些捷径的调查方法都会失去其进行推论的基本要求,根据这些方法收集到的数据,推论的结果是不可靠的”(卢淑华, 1989: 274—275)。董文中 107 个城市的 116 个最低生活保障线的数据资料并非通过随机抽样所得,对其进行回归分析只能用来说明现有数据之间的依存关系。不论其回归模型的拟合程度如何,对总体来说都没有推论价值。既不能使用该模型去预测其他城市的最低生活保障线,也不能用它来预测未来的情况。因此,对其进行假设检验也就毫无意义。而董文中的很大一部分内容就是对文中给出的总量指标模型、平均指标模型、百分比模型和综合指标模型进行了逐一的假设检验,而且检验的显著性水平都达到了 $p < 0.001$ 。如果仅从检验结果看,总体中每对变量之间的线性关系是很显著的。但由于不符合随机抽样的前提条件,所以这些检验本身是没有任何价值的。总体中的变量之间是否存在文中所述的线性关系是不得而知的。

三、在数据来源具有一定效度和信度基础上的统计分析才具有科学性

统计分析只不过是一种工具,分析方法本身的科学性并不能取代资料收集的信度与效度。从童文的资料来源看,用于回归分析中自变量的数据资料主要来源于1996、1997年的《中国城市统计年鉴》,可以认为这些资料是有一定信度和效度的。但用于因变量的城市居民最低生活保障线标准的数据资料中“大部分来自于民政部救灾救济司1998年2、3月编印的《城市居民最低生活保障制度文献资料汇编(一)(二)》;小部分来自笔者参加1999年4月民政部在武汉召开的‘全国城市居民最低生活保障工作管理人员第一期培训班’时,通过询问各省市的基层民政工作人员,以及报章杂志的零星摘录”(童星、刘松涛,2000)。也就是说,这些城市居民最低生活保障线是由各省、市的民政部门“按照当地维持城市居民基本生活所必需的衣、食、住费用,并适当考虑水电燃煤(燃气)费用以及未成年人的义务教育费用确定”(《城市居民最低生活保障条例》,1999)的。而对于这些标准的制定,童文中有这样的论述:“现在许多城市在制定居民贫困标准线时,基本是在本身财力所能允许的范围内‘左顾右盼’,相互参照后制定的,主观随意性很大。”也就是说这些数据制定得是否科学、合理还有待商榷。然而,童文却用这些“主观随意性很大”的数据进行回归分析,由此所获得的数学模型是否能真实、科学地反映城市居民的最低生活保障线,则值得怀疑。“正确的分析来源于正确的收集资料,因此,如果数据的收集没有按照统计分析的要求去做,或者资料的收集不可靠,那么统计分析的结果就可能是虚假的”(卢淑华,1989:18)。

四、统计分析只是手段而不是目的

阅读《城市居民最低生活保障线的测定》一文以后,感到其研究目的似乎并不明确。从文中看,其研究目的是“从少数几项甚至仅仅一项可以从《统计年鉴》等统计资料中顺手拈来的数据,推算出某个城市居民最低生活保障标准线”。事实上,城市居民最低生活保障标准线的确定方法并不复杂。只要具有了诸如城镇居民最低饮食费用或家庭月平均收入等基本数据,用恩格尔系数法或用国际贫穷标准线法很容易算出,关于这一点童文中已有阐述。而童文中所给出的利用回归分析预测最低生活保障线的方法既不如国际贫穷标准线那样简单、易行,也不如恩格尔系数法和维持生活基本费用法直观和具体。所使用的基础数据的来源也不比原来的方法容易。这实质上是使用复杂的方法去解决一个完全可以用简单方法解决的问题。除此之外,童文中还存在着其他将简单问题复杂化的现象。例如,童文中用了很大的篇幅做了没必要做的多重共线性检验。因为“当自变量之间高度相关时,回归方程中的自变量就会互相削弱各自对Y的边际影响,使本身的回归系数的数值下降而其标准误扩大,于是就会出现回归方程整体显著,但各个自变量都不显著的现象,这种现象就称为多重共线性”(郭志刚,1999:49)。童文中的总量指标多元回归方程的F值为61.874,三个回归系数的t检验值分别为5.326、8.328和3.594;平均指标多元回归方程的F值为65.285,三个回归系数的t检验值分别为5.603、8.901和5.128(见童文表5),在大样本的情况下,这些数值足以说明两个多元回归方程不仅整体非常显著,而且每个自变量也都非常显著,是不存在多重共线性问题的。即使为了保险起见进行一下检验的话,只用一个指标就可以说明的问题,童文却用多个指标重复说明。尤其是容许度和方差膨胀因子两个数值互为倒数,仅为检验共线性是否存在时,完全没有必要同时使用。再如:在建立回归模型时,只需列出几个重要参数的值和假设检验结果即可,但童

文中却将许多并非必要的数字罗列出来。看似作了很多统计分析工作,实则使文章的赘信息量大大增加。众所周知,科学方法的应用应注重用简单的方法解决复杂的问题,用简单的理论说明复杂的现象。如果反其道而行之就背离了科学的主题,不是把科学方法的应用作为手段而是作为目的。

五、对回归模型的诊断,重要的是结果而不是过程

多元线性回归模型的建立和使用需要满足一定的基本假设。这些基本假设包括:1. 自变量之间不相关,即没有多重共线性;2. 随机误差项即残差项具有零均值和同方差;3. 随机误差项在不同样本点之间是独立的,不存在序列相关等等。因此,对于使用样本值建立的回归模型,除了进行拟合优度和显著性检验外,还应对模型是否满足假设条件进行检验。由于上述检验的复杂运算过程是由计算机完成的,对于研究者来说,重要的是对计算机输出的结果进行比较、评价和解释,以确定模型是否满足基本的假设要求。童文中用大量的篇幅描述了拟和优度、共线性诊断和残差分析等的原理、过程,列出了大量无太大价值的数据和图表,却对用于残差分析的 Durbin-Watson 检验结果作了错误的解释。童文中是这样阐述其检验结果的:“从表9中得知,该模型的 Durbin-Watson 统计量 D 值为 1.116,说明相邻两点的残差呈正相关。”D 值落在正相关区间,可见残差检验没有通过,本该到此止步。然而,童文却对这一结果作出了错误的分析,认为“原因在于三个自变量国内生产总值、地区分布和城市级别在时间关系上是先后承继的。这使得该模型在残差的独立性上出了一些问题”。笔者认为,童文中的三个自变量根本没有时间上的先后承继关系。存在序列相关的原因可能有两点:一是童文将截面数据与时序数据(9个城市使用了两个年度的数据)混合使用,而时序数据是最容易出现序列相关的。二是可能有些重要的解释变量没有被引入模型所致。无论如何,既然残差的独立性假设检验没有通过,这便在相当程度上决定了模型的有效性,如果不进行修正,那么在此基础上作出的预测也必然无效(李子奈,1992)。

总之,社会统计学应用的目的是为了对事物进行更深入的研究和认识,但必须保证其应用的前提条件得到满足。童文中把不具备推广价值的结论推向总体,使用无效模型进行预测,出现了前后自相矛盾的结果。任何一种方法的应用都有其局限性,如果超过它的使用范围来应用它,或者用不正确的方法来使用它,就可以得出错误的结论。因此,应该用科学的态度来对待科学方法的应用。

参考文献:

- 《城市居民最低生活保障条例》,1999年9月28日中华人民共和国国务院令第271号。
- 郭志刚主编,1999,《社会统计分析方法——SPSS 软件应用》,中国人民大学出版社。
- 何晓群,1999,《现代统计分析方法与应用》,中国人民大学出版社。
- 李子奈,1992,《计量经济学》,清华大学出版社。
- 卢纹岱,2000,《SPSS For Windows 统计分析》,电子工业出版社。
- 卢淑华,1989,《社会统计学》,北京大学出版社。
- 陆璇,1999,《应用统计》,清华大学出版社。
- 童星、刘松涛,2000,《城市居民最低生活保障线的测定》,《社会学研究》第4期。

作者系哈尔滨工业大学哲学社会学系副教授,硕士
责任编辑:张志敏