

也谈科学的方法应予以科学的应用^{*}

——与园田茂人、张汝立商榷

徐道稳

近年来,我国社会学的定量研究在深度和广度上都取得了长足的进步,定量研究方法的使用也呈多元化的趋势。在众多的定量研究中,绝大多数是富有建设性的。它们为社会学研究的规范化和精确化,对定量研究方法的普及都做出了应有的贡献。同时,少数定量研究也存在不够规范和为定量而定量的倾向,对此已有学者撰文指出(尹海洁,2000)。本文则就园田茂人和张汝立在《社会学研究》2000年第1期上发表的《职业评价的中日比较》(下称园田文)在定量研究过程中存在的问题与两位学者商榷。

比较研究的前提是研究对象要具有可比性

比较研究有纵向比较和横向比较之分。纵向比较是历史的比较,即比较同一事物在不同时间内的具体变化;横向比较是不同的具体事物在同一标准下的比较,目的是寻找异同并探索其原因。无论是哪一种比较,首先要确认的是研究对象要具有可比性,即属于同一种类、同一条件或同一关系。

园田文显然属于横向比较研究,但这个研究缺乏可比性。园田文注意到哈尔滨市调查与日本的SSM(Social Stratification and Social Mobility)调查的不可比性:“一是哈尔滨市调查属于市一级水平,而SSM调查属于国家水平;二是这两项调查所使用的抽样方法不同”(园田茂人、张汝立,2000:113)。作者对第一个问题的解决方法是,用哈尔滨市作为中国城市的代表进行分析。至于能不能代表、代表性的好坏则“不应过分强调”。对第二个问题的解决方法是,两个样本虽然无法直接进行比较,但可以选择部分相似的项目进行比较。结果作者“选取两国调查中相似的4种职业,即行政机关的官员、农民、大工厂的工人和个体户来进行比较分析,它们正好代表了传统的士、农、工、商职业结构”(园田茂人、张汝立,2000:114)。姑且不论这4种职业能否反映现代社会的职业结构,仅看它们在两国调查中能否比较。作者的意思是两个样本在整体上不能比较,但是作为部分的4种职业是可以比较的。然而,这种思路在逻辑上是没有说服力的,实际上也是行不通的。请看哈尔滨市的样本构成。哈尔滨市的调查采用定额抽样法,按11个职业类别各抽取100人(从园田文的注释看,调查总数似为1200人),其中高三学生100人,理科大学生100人,文科大学生100人,三类学生共300人,占样本容量的1/4,显然,该样本是非随机的,且样本构成严重偏态。哈尔滨市的这种非随机且严重偏态的样本,不但决定了中日两国的资料在总体上不可比较,而且决定了部分资料也不可比较。在这里,作者

* 本文的写作得到许欣欣博士的指点,谨表谢忱。文中错误由作者自负。

是明知不可为而为之,这种情况在园田文中的其他地方还有体现。例如,作者在表2中列出台湾地区对有关职业的评分,但台湾地区资料的赋值方法与哈尔滨市调查不同,样本构成也有差异,“所以并不一定能进行正确比较”;况且,台湾资料中有6个职业的得分是空白,因此表2中列出台湾地区的职业评分意义不大,属累赘信息。

非随机样本资料一般用作探索性研究,若用作验证理论假设就会得出不准确甚至错误的结论。著名统计学家费歇尔(R. Fisher)在强调选样的随机性时指出,随机性有两个优点,一是随机选择不会偏袒任何因素,避免研究者不自觉的偏见;二是能使调查结果进入数学研究,大数定理为数学研究提供了保证(转自刘大椿,1985:101)。园田文只注重分析本身,而不注重样本的随机性这一分析的前提,这使其后面的分析解释工作陷入困境。根据哈尔滨调查资料建立的多元回归模型和因子分析模型与数据拟合不好就是明证。园田文把每种职业声望得分与7个评价职业的标准(重视因素)做了多元回归,结果多元回归模型对日本资料拟合较好,而对中国的拟合则不理想。为什么中国的职业声望得分与重视因素之间拟合不好呢?作者认为,这是中日两国不同职业群体对职业评价的差异造成的。为此,作者以官员、农民、工人和个体户4种职业为例对中日两国作了比较。笔者对这种解释表示怀疑。首先,中国的样本是非随机的且构成很不合理,由此得出的回归模型拟合效果不好正在情理之中。其次,中国的职业得分是由11类职业群体评价中得出的结果,即使真的是由不同职业群体对职业评价的差异造成拟合效果不好,那么,仅用作者选的4种职业群体的评价是否真的能解释也值得怀疑。

统计分析方法不能盲目选用

每个定量研究都要用到一种或几种统计分析方法,但具体运用哪些方法,则需要研究者根据研究的目的、调查资料的特点以及各研究方法的适用范围进行选择。例如:“因子分析的目的在于简化数据或者找出基本的数据结构,使用因子分析前提条件是观测变量之间应该有较强的相关关系。如果变量之间的相关程度很小的话,它们不可能共享公因子”(郭志刚,1999:93)。

园田文的主要统计分析工具是因子分析和多元回归分析,但从分析的过程和结果看,园田文对这些方法的应用显得有些盲目。首先,园田文没有正确和有效地利用相关矩阵中的信息。在做因子分析之前应先算出相关矩阵,然后对相关矩阵进行检验,看是否适合使用因子分析方法。如果相关矩阵中大部分相关系数都小于0.3,则不宜做因子分析。对于相关矩阵的检验,SPSS软件提供了三种检验方法,即反映象相关矩阵(Anti-image correlation matrix)检验,巴特利特球体检验(Bartlett test of sphericity)和KMO(Kaiser-Meyer-Olkin Measure of Sampling Adequacy)测度(郭志刚,1999:93)。园田文在做职业评价的结构特征分析时没有做相关矩阵分析,在做评价标准的结构特征分析时却在因子分析之后列出了相关矩阵(见园田文表5),目的是“加深理解”,“进一步检验上述分析的可行性”(园田茂人、张汝立,2000:118),这实在是本末倒置。

另外,园田文表5列出了所有的相关系数和显著水平,令人眼花缭乱,显得很不规范。实际上,相关矩阵中的相关系数关于对角线对称,只要列出对角线一侧的相关系数且在检验显著的相关系数后以星号标注即可。为此,本文对其修改后作为表1列出。从表1中不难看出,在日本的21个相关系数中有11个相关系数小于0.3,其他10个系数也不大;而在中国的21个

相关系数中则有 16 个系数小于 0.3, 其他 5 个系数也不超过 0.4。这说明大部分相关系数太小, 园田文中的调查资料不适合做因子分析。

其次, 园田文对因子命名和因子负荷认识模糊。作者在文中对 20 种职业声望得分进行因子分析(见园田文表 3)后指出, “令人感兴趣的是, 中日双方都抽取了‘职业无贵贱’因子, 这一因子在日本为第二因子, 在中国则为第一因子, 其顺序逆转。当然, 无法判断这一因子是否真实, 但考虑到中国长时期所进行的‘职业无贵贱’的教育效果, 这一差异是可以理解的”(园田茂

表 1 重视因素得分的相关矩阵: 中日比较

	教育	技能	收入	尊敬	贡献	影响
技能	0.3377 *** 0.2094 ***					
收入	0.3964 *** 0.1092 *	0.2180 *** 0.0846				
尊敬	0.4064 *** 0.0925	0.2716 *** 0.1618 ***	0.4081 *** 0.0464			
贡献	0.2167 *** 0.1770 ***	0.2804 *** 0.3885 ***	0.1427 * 0.0744	0.5031 *** 0.2595 ***		
影响	0.2614 *** 0.1182 *	0.2386 *** 0.2745 ***	0.1494 * 0.0845	0.4796 *** 0.2651 ***	0.6283 *** 0.3545 ***	
创造力	0.1222 * 0.0942	0.3348 *** 0.3615 ***	0.0622 0.0289	0.2625 *** 0.1858 ***	0.3910 *** 0.3523 ***	0.3970 *** 0.3993 ***

注: 每一项的第一行数值表示日本的相关系数, 第二行表示中国的相关系数。

* $p < 0.05$; *** $p < 0.001$ 。

人、张汝立, 2000: 115)。作者没有说明得出“职业无贵贱”因子的依据, 但从行文中似乎可以看出他们的依据是, 日本的第二因子负荷和中国的第一因子负荷全是正数。如果真是这样的话, 就是对因子负荷的误解, 也是对因子命名的误解。这样的误解在园田文关于表 4 数据的解释中再次出现(园田文的表 4 即本文的表 2)。根据表 4, 日本的第一因子中, 所有的因子负荷都是正值, 作者给该因子命名为“都重视因子”; 第二因子中, 因子负荷有正有负, 作者没有给它命名。中国第一因子中, 收入标准的因子负荷是-0.060, 作者可能据此认为, 收入标准的因子负荷是负数就说明了调查对象对金钱不重视, 因此给第一因子命名为“非金钱评价”因子, 或“为人民服务”因子, 第二因子为“金钱评价”因子。这些因子的命名令人啼笑皆非。实际上因子的得名和因子的重要程度与因子负荷的正负号没有关系, 而与因子负荷的绝对值有关(柯惠新等, 1992: 371, 487)。因为, 因子负荷的正负号只决定变量在因子坐标中的位置, 并不影响变量与因子之间的关系。此外, 因子负荷的正负与因子坐标的转换有关, 在一个因子坐标中为正, 在另一个因子坐标中可能为负, 反之亦然。所以, 不能用因子负荷的正负号来判断因子的重要性, 或者给因子命名。关于因子的命名, 通常在因子旋转之后进行。提取因子后, 一般要进行因子旋转, 通过坐标变换使因子解的意义更加明确, 故因子的得名一般在因子旋转后方才变得清晰。

第三, 园田文中的多元回归模型很不规范。众所周知, 根据调查资料建立的回归模型在多大程度上解释了因变量的变化, 对调查资料的拟合程度如何, 要用决定系数 R^2 来衡量。在多元线性回归分析中, 决定系数越大, 说明回归方程拟合数据越好, 即回归方程中的自变量对因

变量的解释力越强;如果决定系数过小,表示在所建立的方程中很可能漏掉了某些重要的因素(郭志刚, 1999: 35;袁方、王汉生, 1997: 561)。回归系数是回归分析中的又一重要参数。某一自变量的回归系数,表示在其他自变量保持不变时该自变量与因变量的变化关系。回归系数的大小与变量所取的单位有关,故不同变量的回归系数之间通常不能进行比较,但可以通过变量的标准化消除自变量的测量单位,变回归系数为标准回归系数,从而使各自变量对因变量影响的大小变得可以进行比较。在多元回归分析中,一般既列出回归系数,又列出标准回归系数。显著性检验是回归分析的另一重要内容。检验包括两个方面,一是回归方程的检验(F检验),如果检验显著,说明回归方程对总体是有效的;二是回归系数的检验(t检验),如果检验显著,可以认为总体中自变量和因变量之间存在显著的线性关系。综上所述,决定系数、回归系数(包括标准回归系数)、显著水平是衡量回归模型的三个重要指标。但令人费解的是,园田文在把中日两国的20种职业声望得分,分别与7个评价标准做的多元回归分析(见园田文表6、7)中,只标明显著水平,而没有给出回归系数和决定系数,这就使读者无法鉴别回归方程的有效性,也无法比较各自变量对因变量的主次作用。在此需要特别指出的是,园田文表6中共有20个回归方程,其中6个方程中所有自变量无一显著,这标志着这些多元回归模型存在严重的多重共线性问题^①。但园田文对此却视而不见,居然利用这一模型继续讨论并得出结论。

综上所述,不难看出,园田文不仅在因子分析方面的使用上随意性很大,而且在多元回归方法的使用上亦颇不得要领。

对分析结果的讨论不能只见树木不见森林

对分析结果的讨论是定量研究中的最后一步。这一步要求研究者根据分析的结果作出客观全面的判断,或验证假设(部分验证),或推翻假设(修正假设)。园田文由于定量研究方法应用上的失当,导致其得出的结论牵强附会,漏洞百出。限于篇幅,这里仅举一例。

园田文在对评价标准的结构特征作了因子分析(见园田文表4)后得出结论:中日两国在收入与学历的关连性以及收入与创造力的关连性上均逆转。对此作者解释说:“它表明在日本的特征是高学历高收入,而在中国则是高学历低收入”;“在日本,一种职业能否获得高收入,并不在于它发挥创造性的程度,而在中国则是能发挥创造性的职业获得高收入”(园田茂人、张汝立, 2000: 118)。为了看清上述结论的武断性,现把园田文表4作为本文表2列出。

表2 评价标准的结构特征: 因子分析的中日比较

	日本		中国	
	第一因子	第二因子	第一因子	第二因子
教育	0.575	0.536	0.358	0.470
技能	0.556	0.070	0.670	0.167
收入	0.468	0.682	-0.060	0.884
尊敬	0.773	0.139	0.492	0.218
贡献	0.758	-0.350	0.725	-0.039
影响力	0.752	-0.334	0.687	-0.029
创造力	0.570	-0.450	0.692	0.218
贡献率	41.7%	17.3%	32.8%	15.5%

① 关于多重共线性及其解决方法,参看郭志刚主编的《社会统计分析方法——SPSS软件应用》中的有关章节。

在上表中,日本的第一因子中,教育的因子负荷为正,收入的因子负荷也为正;而在中国的第一因子中,教育的因子负荷为正,收入的因子负荷为负,作者据此得出“在日本的特征是高学历高收入,而在中国是高学历低收入”。在日本第二因子中,创造力的因子负荷为负,收入的因子负荷为正;而中国的第二因子中,创造力的因子负荷为正,收入的因子负荷也为正,故“在日本,一种职业能否获得高收入,并不在于它发挥创造性的程度,而在中国则是能发挥创造性的职业获得高收入”。作者可能忘了,表2中教育、收入等7个变量是调查对象对职业评价时所用的7个标准,并不能反映调查对象的教育、收入等变量之间的相关关系。作者显然是把职业评价中的教育标准与收入标准的关系,误认为是教育与收入本身的关系;把创造力标准与收入标准的关系,误认为是创造力与收入本身的关系。可见,作者的前述结论具有很大的盲目性和主观随意性。

园田文中出现的问题再次表明,统计分析仅仅是一种工具,只有在正确的方法论指导下,“只有当对现象的本质和内部联系有了一定的认识之后,才能考虑某种确定的统计工具。否则,轻率地选择某种统计工具,或者干脆靠计算机输出来代替分析,所得的结论都是不可靠的,甚至是危险的”(袁方、王汉生,1997:561)。

参考文献:

- 冯士雍、施锡铨,1996,《抽样调查——理论、方法与实践》,上海科学技术出版社。
- 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。
- 柯惠新等,1992,《调查研究中的统计分析法》,北京广播学院出版社。
- 刘大椿,1985,《科学活动论》,人民出版社。
- 卢淑华,1997,《社会统计学》,北京大学出版社。
- 尹海洁,2000,《科学的方法应予以科学的应用——与童星、刘松涛商榷》,《社会学研究》第6期。
- 袁方主编、王汉生副主编,1997,《社会研究方法教程》,北京大学出版社。
- 园田茂人、张汝立,2000,《职业评价的中日比较——SSM 调查与哈尔滨市调查的比较分析》,《社会学研究》第1期。

作者系深圳大学法学院社会学系讲师,硕士
责任编辑:张宛丽