

## 从类型逻辑和总体逻辑理解定量社会学新范式

周旅军 北京大学社会学系

面对纷繁复杂、变动不居的社会现实，我们无法在严格意义上重复同一研究，人为干预几乎不可能或效果微不足道，最重要的是记载和理解现实中人口的实证形态。量化方法不应成为发掘普遍规律，从而描绘和解释所有个体行为的工具。

在当代社会科学研究的后实证范式中，研究方法以量化分析为主，其主要依据为自然科学方法论，同时承认社会现象与自然现象的本质区别、承认社会理论的历史性、相对性。谢宇从科学哲学视角指出，存在类型逻辑思维和总体逻辑思维两种传统，代表着社会科学与自然科学在本体论、认识论以及方法论上的差异。

变异是真实的，并非失误

在自然科学中，只要我们理解典型现象，就可以将其概括并推广到个体和具体问题上。变异则是对本质世界的拙劣复制，通过抽象思维将偏差(deviation)忽略之后才能得到真正的知识。科学家只有懂得了如何超越由形成的世界产生的偏差干扰，才能探求到伟大的科学发现。

社会科学研究的是“形成的世界”，现实总是在不断变化。对此，凯特勒(Lambert Adolphe Jacques Quetelet)的社会物理学认为，要在社会科学中获得真理就应该取平均值。尽管个体的行为各异、无法预测，总体或子群的平均值却相对稳定可测。查尔斯·达尔文提出总体逻辑思维(population thinking)，对类型逻辑思维提出挑战和修正，其后由弗朗西斯·加尔顿引入社会科学。他认为变异是真实的，并非失误，进化的前提就是个体间的差异性。相反，平均数只是总体的特征值之一，是探讨真实原因的手段，而非原因本身。社会科学的重要性在于研究个体之间何以存在差异，不能仅比较平均数，侧重点应集中在变异(variation)和共变(covariation)上，关注“属性是如何分布的”。加尔顿对个体差异和变异而非平均值的重视最终使他发现了“回归”和“相关”的概念。

在总体科学(比如社会科学)中，观测到的各个体的取值不同，其原因不仅来自于测量误差，还因为它们是一总体中本质上不同的个体。研究重点是用最简洁的描述概括总体差异。尽管估算方法与自然科学相同，但研究目标、隐含的假设和对结果的解释存在本质差别。面对纷繁复杂、变动不居的社会现实，我们无法在严格意义上重复同一研究，人为干预几乎不可能或效果微不足道，因此，统计分析只能“旨在强调如何解释结果，而非提供寻找原因的方法”，最重要的是记载和理解现实中人口的实证形态。量化方法不应成为发掘普遍规律，从而描绘和解释所有个体行为的工具。

社会科学重视两类差异

我们可以通过对统计模型如多层线性模型在社会科学研究中的必要性和局限性的反思来理解方法论上的争论。社会科学重视差异，包括个体与个体之间的差异和由个体组成的整体之间的差异。社会现象存在差异首先就表现在分析单位的不同层次上。研究中出现的生态学谬误(ecological fallacy)错在把高层次的信息、经验应用到低层次的分析单位上，将从整体数据中得出的模式不加反思地强加于个体，混淆了不同层次主体的行为模式。另外，群体变异的模式会随社会情境的变化而变化。社会情境通常由时间和空间界定，它不同于社会分组，是一个独立的社会系统。个体变异的模式可以被宏观条件，如社会结构等支配。因此将宏观分析和微观分析结合起来，在研究框架中同时包括不同层次的分析单位成为研究者努力探索的方向。

具体而言，社会科学研究中的数据结构常常体现为分层嵌套的形式，即低一层的数据嵌套于

高一层的结构之中，比如居民嵌套于社区。研究的问题涉及：1.哪些个体水平的解释变量影响结果变量？2.哪些情境变量影响结果变量？3.个体水平解释变量和结果变量之间的关系是否随个体所处的社会情境的特征变化而变化？4.哪些情境变量调节个体水平解释变量的效应？

针对多层数据，传统的统计技术主要有两种分析路径，一种是在低水平层次上进行回归分析，忽略高水平层次上的差异；另一种是把低水平变量整合到高水平层次上，忽略个体水平的特征，丢弃所有的组内变异，以致无法达到解释变异的目的。同时，常规单一水平的统计模型，如普通线性回归、方差分析等，其运用前提要求个体间的随机误差独立，然而取自同一总体的个体与随机抽样相比，具有更高的相关性，假设无法得到满足。此外，在纵向调查数据的重复测量中，对于个体水平上的多次测量，由于个体特征相同以及测量间的相互影响，也存在测量误差之间的自相关情况。

为此，需要使用多层性模型，其关键在于分析环境层次和个人层次的自变量的交互作用，同时考察微观和宏观变量的效应。与传统的回归分析相比，多层线性模型可以通过计算不同水平变异在总变异中所占的比率来确定不同水平对因变量的影响程度。过去只强调个人因素及其构成特征的简化论观点和只强调环境和组织因素的生态学观点的争执，有很大局限。

简化引起的误差是理解上的缺陷

在总体逻辑思维看来，社会科学的变异性不仅指个体间的变异，也包括同一个体在成长过程中的变异。也就是说，除了与生俱来的变异性，还有个体在发展中产生的变异性，它受不确定因素影响，不一定可以遗传，也不一定有物质性载体。当对相同的观测对象进行重复测量时，可以将所得数据本身看成嵌套结构，进行多层线性模型分析。这样不仅可以分析总体上个体随时间产生的变化，还能将个体之间增长的差异进行分析，对原因作出解释。

多层线性模型的局限性在于所使用的参数估计方法较传统估计参数的方法要复杂得多，而且与潜变量增长曲线模型相比，它不能处理变量之间间接的影响关系，也不能处理复杂的观测变量和潜变量之间的关系，只能分析变量之间的直接因果路径，对潜变量之间关系的分析要比潜变量增长曲线模型复杂，在测量模型上也有更多的限定条件。

事实上，除了特定的技术细节，定量分析方法在实际应用中所具有的局限性也通过多层线性模型等统计模型反映出来。虽然社会科学研究只能提供关于总体的少量参数，这看上去似乎与柏拉图的想法一样，但仅是为了理解和讨论的方便。简化(parsimony)的代价体现为误差，它不是自然科学中的测量误差，并不能够忽略不计，因为它是一个真实的、由于不充分的知识而导致的理解上的缺陷。

正是通过对类型逻辑思维和总体逻辑思维的区分，我们才能更好地理解 Otis Dudley Duncan 提出的定量社会学新范式的核心原则，即强调实证现实的首要性，反对模仿自然科学去寻求社会普遍规律的做法。定量工具不应用于发现所谓能够描述或解释所有个体的行为的普遍规律；相反，它应该概括组间(between-group)差异的实证模式，并暂时忽视组内(within-group)的个体差异。随着逐步在分析中增加复杂性，社会科学研究者能够提升对充满总体异质性的世界的理解。