

定量统计的方法和思想能够让考古学意识到什么问题？

——对《统计分析在商代陶鬲研究中的应用》一文的 批判与再研究*

陈锦航*

摘要：本文首先对《统计分析在商代陶鬲研究中的应用》一文进行了批判，揭示了该文由于不理解定量统计思想而导致统计方法运用不当；其后使用该文同样的数据，基于陶鬲形态，分别对商文化二期说、三期说进行验证，发现二期说比三期说在相关系数、显著水平、以及参数检验指标上更有解释力；最后从定量统计思想出发，重新反思了考古界中样本和总体关系、考古学分期的中定性和定量的关系等议题。

关键词：统计思想 定性与定量 商代陶鬲 考古分期

随着我国考古学的研究的深入、计算机技术的普及和推进，在科技考古、考古文化圈的判定、器型分析、墓葬大小与随葬品关系等分析上均有数学分析的运用，此类运用数学定量研究方法引入考古学被我国考古界视为“定量考古”交叉学科的确立和兴起^[1]。广义而言，凡是运用到数学知识进行考古研究均可视为定量考古^[2]，在具体实践领域中统计学则运用较多。

遗憾的是，由于我国很多考古学家缺乏专业统计学的系统训练和学习，在没有深入理解统计学思想及研究方法论就使用统计分析，导致实际研究多停留在模仿阶段、将软件结果输出作为辅助论证，甚至出现一些误用和滥用的情况，相关研究论文评审专家由于缺乏相应知识，一些统计分析不合格的研究也得以发表，对于学界造成了负面示范作用。本文以《统计分析在商代陶鬲研究中的应用》^[3]一文为批判分析对象，通过指出该文统计误用和滥用的一些情况，并采用相对正确的统计方法重新对该文的相关数据进行分析，这也是国内首次一个纯社会学出身的研究者利用考古二手资料，从交叉学科的角度反思定量考古方法论的一篇研究，希望以实际的论文批判与再研究推进考古学对定量研究的反思和运用。

一、对《统计分析在商代陶鬲研究中的应用》一文中定量思想与统计方法不当之处的分析

（一）定量研究思想认识上的错误

该文框架完全没有按照科学假设检验思路展开，既然存在两种分期命题，就需要对三期说与二期说分别进行操作化验证，而该文仅仅在认可三期说的前提下，用定量方法进行事后补充论证。此外该文完全没有统计推论检验的思路，即没有样本（153 个陶鬲）——总体（商代陶鬲）的区分认识，但是又默认样本陶鬲就可以代表商代陶鬲。下面就该文操作化与统计分析错误之处进行详尽批判。

（二）统计方法错误

* 本文在初稿成文后于 2021 年向驻周原考古基地的师生进行了汇报，得到了社科院考古所宋江宁老师及原宝鸡考古所辛宇宝贵修改意见，特此致谢。为了满足其它研究者对数据进行重复检验，本文涉及的相关研究数据可通过百度网盘下载（网址链接：<https://pan.baidu.com/s/11GltFWCd3E1jQaP5mA5EGA> 提取码：0525,SPSS.SAV 文件）。

* 作者简介：陈锦航，男，纯社会学学科背景，2017 年获北京大学社会学博士学位，原法门寺博物馆副研究员，现就职于兰州文理学院。本文删节版分两部分发表于《科学与财富》2023 年 15 卷中、下。

1.数据标准化处理没有抓住重点

由于样本中殷墓出土的陶鬲是明器，体积较小，无法直接与其它非明器样本进行比较，确实需要进行标准化处理，使得单位级别一致，该文的标准化转置确实可以解决单位级别问题，但是这和研究关注点——陶鬲形态没有任何关系，在下面的研究路径一种，将用 最大腹径与通高比值 这一变量直指研究核心问题。

2.因子分析仅仅是降维数据整理技术

因子分析是将多个实测变量转换为少数几个综合指标（或称潜变量），它反映的是一种降维的思想。通过降维将相关性高的变量聚在一起（最大腹径、通高、档高），从而减少需要分析的变量的数量，而减少问题分析的复杂性。该研究因子分析结果没有问题，但是该文在第三部分开头对于因子分析的介绍完全没有理解因子分析的目的，就是说因子分析仅仅是降维数据整理技术，而不是多元统计分析方法。这三个变量合成一个因子以后它与所要解决的研究议题有什么关系？操作性逻辑何在？这个因子应该如何命名？

3.为了直观看散点图，因子分析是“必要的”

正是由于进行了因子分析，将四个变量整合为两个因子，才可以画出散点图，可以通过肉眼观察散点图下结论，由于我们所处的世界就是一个三维空间，超过三个维度的数据是没办法用几何方式画出来的，也就是说，四列以上的数据就无法通过统计图呈现了。

如果不做因子分析，完全可以用多元方差分析进行更有效地统计分析，四个变量对于多元方差分析并不算多，统计软件也会提供统计检验数据。可以推断，该文作者并不真正理解多元统计分析技术，很可能仅仅是由于前人在类似研究中用过因子分析（该文开头两段提到），才盲目模仿这一方法。其深层次的原因应是定量考古对于统计图具有迷信般的偏好，特别是散点图、柱状图等，图确实可以直观地、形象地呈现结果，但是图的信息远没有统计表与精确，也没有细节信息（如方差、显著度水平），比如观察散点图好像分布集中一点就可以说明呈现相关或拟合，而图之后完全没有相关系数和统计检验数据的报告（此问题在今天考古界研究中依旧存在），这完全是直观印象，是对于统计数据不敏感的表现，更是不专业体现。虽然有一定可能通过观察统计图做出正确的判断，但是其实可以仅用相关系数值、分布情况值就能够精确、到位地做出准确地判断。

需要提醒各位读者注意的是，导致上面统计技术性运用出现重大失误的根本性原因是该文对定量考古两大思想性错误认识，在本文将在结论部分进行深度反思。

（四）该文统计呈现细节不规范

1.统计软件的选择无需提及

该文提到“分析过程以 SPSS.11 软件实现”，这又体现了作为定量论文的不规范之处，只有非专业定量研究者才刻意强调所使用的统计软件。统计分析最为核心的是要说明统计方法运用是满足数据条件的，能够回答所要解决的问题。使用的统计软件是完全无需提及的，统计软件有很多，而且也在持续版本更新，只要能够进行需要的统计分析，用什么软件，哪个版本均不重要，从来就没有因为统计软件版本更新、功能更强大，统计软件就更正确，即便人工进行统计分析也是可以的。

2.统计分析结果展现不应直接从统计软件粘贴过来

该文的散点图、统计结果表均直接从统计软件中粘贴而来，这亦是不规范之处。定量论文中只应呈现所必须的统计结果（可参照心理学、社会学定量论文作为规范），一些统计结果表格可以从一张表合并呈现就应在一张表上呈现，而不是多个表。且其中的英文应当翻译为中文。

3.样本数据信息无需完全展示

该文表一是 153 个样本数据呈现，文中提到的“定量描述”就是指该表，占了整整一页。这又是该文定量研究不规范之处，在社科领域，样本量，变量数量是非常庞大的，一般少则

几百个样本数据，每个样本又有几十个变量，是不可能在研究论文甚至专著中展示的，所以仅通过描述统计，如本文的表一，进行资料初步整合式呈现。所以该文的表一完全不是定量描述，而仅仅是原始数据信息呈现。

此外还有一处笔误，变量表所展示的一个变量为“最大腹径”，但是到因子分析和散点图处变为了“最大外径”，结合该文数据可以推断实际为同一个数据，这两个概念指涉是非常不同的，但是对于该文研究结果影响不大。

（五）小结：该文所反映出考古界对定量考古认识两大根本思想性偏误

1. 量化的研究可以代替定性分析过程

该文在摘要中有如下表述：“统计分析方法应用于考古学是以量化的研究代替定性分析过程。”可谓大谬也！如果没有考古对于商代陶鬲所呈现出：瘦体鬲→方体鬲→扁体鬲的定性认识、没有商文化三期说的假说、没有把陶鬲视为商文化研究的重要标志性器物、定量研究是无法开展的，是无法进行操作化的。这也是为何首先需要掌握社科研究方法论再学习统计方法。统计属于定量研究，定量研究的前提和基础是定性研究，定量研究的具体开展需要定性的理解和假设的支持^[4]。

举个考古界众所周知的例子，三代考古多用古陶器组合的比例作为分期与考古学文化类型的标准，不同学者对于核心器物组合的认识是不同的，哪类陶器作为核心标志、标准也是有不同观点、对于陶器到底应当注重器型、还是花纹、还是尺寸？对于青铜器而言，考古界认为器型和铭文比尺寸更为重要，所以对于青铜器的尺寸进行定量研究是非常少见的。此类就是定性问题，只有解决定性问题，才能开展操作化，才能编码录入统计软件进行分析，所以不同的定性认识自然会影响统计的结果。对于统计软件而言，不同类的陶器仅仅是不同的变量，不同的数字。统计分析方法仅仅是分析工具，是不能直接解决上述定性问题的。定性并不是绝对正确的观点，而仅仅是假设，统计能够帮助我们检验这些不同的假设是否成立，哪一种观点从统计角度看相对合理，成立的概率更大、解释力更强，是一种概率性思维，而非是否的确定回答。

2. 定量研究比单纯地定性分析更加精确和深刻

该文在开头处提到了夏鼐先生希望将统计分析运用到考古学，特别是器物类型学的研究中，因为定量研究比单纯地定性分析更加精确和深刻。与之类似，孙庆伟也指出这种基于模糊语言对器物描述进而进行分期的方法是不科学、不准确的^[5]。遗憾的是，这应是该文统计分析上的指导思想根本性偏误，所以也就有了具体统计分析方法的误用。

考古学相关的研究数据虽然难以达到统计学所要求的随机数据，但是所有的考古学研究者均期望通过物质遗存理解古代文明。

从统计学而言，已出土、已发现的陶鬲仅仅是样本，考古学研究者预设了这些陶鬲样本能够代表没有发现或没有保存下来的陶鬲，翻译为统计语言就是样本陶鬲能够推断总体陶鬲情况，所以在定量考古中必然应有统计推断检验的分析这一环节，即便样本随机性不足也是必需的。所以即使孙庆伟整理了不同考古文化陶器组合的百分比数据，但是没有进行统计推断检验，看似精确的百分比、统计表所呈现的也仅仅是形式上的“精确和深刻。”也正是基于上述对于统计在考古学运用定位的不当认识，很多国内定量考古研究中完全没有进行假设检验，统计推论的分析过程。

从考古学而言，对于商代陶鬲的研究从来不是为了研究陶鬲本身，而是期望通过陶鬲回答商代分期问题、商代国家范围、文化特点、甚至断代史问题。但是统计学没有办法直接回答该问题，统计学只能告诉我们通过样本所反映出不同类型的差别是否可以推论到总体陶鬲之上。不论本文还是《统计分析商代陶鬲》一文都是基于陶鬲之不同能够反映商不同时期的文化（分期）这一预设之上的。但是这一预设可能是错误的，也许陶鬲类型虽然有不同，但是文化却没有变化，物质文化有显著变化并不意味着改朝换代，历史上这样的案例比比皆是，

如古埃及第十八王朝因法老阿蒙霍特普四世进行的宗教改革，推崇新的太阳神“阿吞”；古罗马叛教者尤里安，均是昙花一现式地改变了文化，如果没有史料则很容易通过考古判断出现了重大地改朝换代，反之亦然，物质文化虽然没有变化但是出现了重大社会变迁，这种情况在现代社会中是非常常见的。

二、研究路径一：采用标准的定量假设检验流程对该文数据进行再研究

表 1 标准的定量假设检验流程在商代陶鬲研究的运用过程

哲学方法论层面	具体在商代陶鬲研究层面
现象	在不同地点和不同地层学中发现了形态不同的陶鬲
对现象进行归纳认识	从考古类型学出发商代陶鬲可以分为不同阶段
基于归纳提出命题与竞争解释	陶鬲不同阶段对应着商文化分期——商文化三期说 VS 商文化二期说
假设（待检验）	用商代陶鬲所呈现出：瘦体鬲→方体鬲→扁体鬲 的初步经验认识，应当在商文化不同分期下体现出统计学意义上差异
定量数据的选择、假设操作化	商代陶鬲样本数据，通过最大腹径与通高比来让研究假设可以被观察；商文化不同分期可与陶鬲出土考古文化
自变量与因变量	用商文化分期推测陶鬲形态，所以自变量为商文化分期、因变量为陶鬲形态
选择合适的统计方法进行统计分析	数据关系类型为：定序（不同分期）——定距（最大腹径与通高比值），采用单因素方差分析
分析结果判定（证伪还是证实假设）	考察指标：显著性、相关系数 η^2 ，参数检验水平
如均证伪是否可以提出新的解释	

对已有研究单纯进行批判是不够的，需要做出更好的研究才具有说服力。下面将采用标准社科定量研究流程，使用该文提供的资料与数据，从命题阶段开始，进行重新研究。需要提醒读者注意的是，按照上表科学的定量研究逻辑，只要前面一个步骤出现问题，后面的结果就会出问题，整个研究的结果就是不成立的。

（一）命题与竞争解释：商文化三期说 VS 商文化二期说

（二）研究预设：

1.商文化是可以分为不同时期的，有内部差异的；

2.陶鬲为商文化研究的重要标志性器物，已发现的洛阳地区陶鬲形态上的变化能够证实或证伪上面的假设，进而判断哪一种分期说更有解释力；

3.不同考古文化圈类型可以划分在商文化某一期之下

考古界习惯用发掘地点进行初步的考古文化圈命名，就商代分期而言，不同时期属于不同地层学阶段，基于不同地层学划分，对于商文化分期与考古文化圈有如下对应关系：

三期说：

早商：南关外期、二里岗下层一到二期、二里岗上层一期

中商：白家庄期、洹北商城、殷墟一期（苗圃一期、大司空一期、人民公园）

晚商：殷墟二至四期（苗圃二期、三期、殷墓随葬品）

二期说：

前期：南关外期、二里岗上层、二里岗下层一到二期、苗圃一期、大司空一期、白家庄期、人民公园期、洹北花园、

后期：殷墟二至四期（苗圃二期、三期、殷墓随葬品）

基于样本数据，需要对预设 3 进行检验，如果每一分期下子集之内的考古文化圈陶鬲形态没有呈现出内在差异，则可以继续开展研究，如果这一预设被证伪，则按照时期对商文化进行分期这一命题是有问题的，应从空间地域维度考虑对商文化进行区分。即纵向时间文化分期 vs 横向空间地域文化分期。

对预设 2 的检验，单因素方差分析发现每一分期下子集之内的考古文化圈陶鬲形态没有呈现出内在差异（均不显著）。

（三）假设与假设之间关系的判定：

假设 1：若三期说成立，不同时期陶鬲形态有统计学意义上的差异；

假设 2：若二期说成立，不同时期陶鬲形态有统计学意义上的差异；

由于二期说与三期说并非非此即彼的关系，三期说是对二期说的进一步划分从逻辑上而言，如果三期说成立、那么二期说亦成立。二者不是竞争解释关系，所以需要有一个标准来判定到底哪一种解释更有解释力。

从科学认识论角度而言，分类更详细，代表知识越进步，与此矛盾的是，从奥卡姆剃刀原理出发，越简单的解释，就是越好的解释，但是更复杂的分期说并不涉及更多解释实体的引入。考古界对这几种分期自然有不同观点，尚处于争议之中。那么能否从统计学角度确立一个判定哪种假说更具有解释力的标准呢？

（四）数据与变量

1. 数据整理与样本评估

多亏了该文表 1 提供的现成数据，方便笔者直接录入统计软件进行再研究。并结合参考文献对于相关陶鬲样本数据进行了重新核对，发现仅有两处笔误，样本 C5T86(4):53 与 C9T124(2):75 误标了比例尺。

此外，在数据录入过程中，对于有比例尺的样本按照比例尺恢复了实际尺寸数据。

整理后，共有 7 个变量：三期说、二期说、原报告分期、最大腹径、通高、档高、实足高、腹径通高比。

表 2 原报告分期与三期说 交互表统计描述结果

			三期说			合计
			一期	二期	三期	
原报告分期	二里岗下层一期	计数	9	0	0	9
		三期说 中的 %	18.4%	.0%	.0%	6.1%
	二里岗下层二期	计数	30	0	0	30
		三期说 中的 %	61.2%	.0%	.0%	20.4%
	二里岗上层一期	计数	10	0	0	10
		三期说 中的 %	20.4%	.0%	.0%	6.8%
白家庄期	计数	0	5	0	5	
	三期说 中的 %	.0%	12.2%	.0%	3.4%	

人民公园	计数	0	15	0	15
	三期说 中的 %	.0%	36.6%	.0%	10.2%
苗圃 3 期	计数	0	0	16	16
	三期说 中的 %	.0%	.0%	28.1%	10.9%
苗圃 1 期	计数	0	9	0	9
	三期说 中的 %	.0%	22.0%	.0%	6.1%
苗圃 2 期	计数	0	0	11	11
	三期说 中的 %	.0%	.0%	19.3%	7.5%
大司空 1 期	计数	0	7	0	7
	三期说 中的 %	.0%	17.1%	.0%	4.8%
殷墓	计数	0	0	22	22
	三期说 中的 %	.0%	.0%	38.6%	15.0%
南关外	计数	0	0	8	8
	三期说 中的 %	.0%	.0%	14.0%	5.4%
洹北花园	计数	0	5	0	5
	三期说 中的 %	.0%	12.2%	.0%	3.4%
合计	计数	49	41	57	147
	三期说 中的 %	100.0%	100.0%	100.0%	100.0%

表 3 原报告分期与二期说 交互表统计描述结果

			二期说		合计
			前期	后期	
原报告分期	二里岗下层一期	计数	9	0	9
		二期说 中的 %	9.2%	.0%	6.1%
	二里岗下层二期	计数	30	0	30
		二期说 中的 %	30.6%	.0%	20.4%
	二里岗上层一期	计数	10	0	10
		二期说 中的 %	10.2%	.0%	6.8%
	白家庄期	计数	5	0	5
		二期说 中的 %	5.1%	.0%	3.4%
	人民公园	计数	15	0	15
		二期说 中的 %	15.3%	.0%	10.2%
	苗圃 3 期	计数	0	16	16
		二期说 中的 %	.0%	32.7%	10.9%
	苗圃 1 期	计数	9	0	9

	二期说 中的 %	9.2%	.0%	6.1%
苗圃 2 期	计数	0	11	11
	二期说 中的 %	.0%	22.4%	7.5%
大司空 1 期	计数	7	0	7
	二期说 中的 %	7.1%	.0%	4.8%
殷墓	计数	0	22	22
	二期说 中的 %	.0%	44.9%	15.0%
南关外	计数	8	0	8
	二期说 中的 %	8.2%	.0%	5.4%
洹北花园	计数	5	0	5
	二期说 中的 %	5.1%	.0%	3.4%
合计	计数	98	49	147
	二期说 中的 %	100.0%	100.0%	100.0%

删除掉明器和没有比例尺的样本，最大腹径、通高、档高、实足高、与包含明器的腹径通高比，五个变量描述结果如下：

表 4 商代陶鬲指标变量描述

	N	极小值	极大值	均值	标准差
最大腹径	137	8.30	31.50	17.7910	3.18067
通高	137	8.96	36.00	17.8642	4.31741
档高	137	.50	12.60	4.8996	2.23875
实足高	136	.50	8.10	3.5014	1.50085
腹径通高比	147(含明器)	.62	1.48	1.0212	.15551

(四) 操作化

瘦体鬲、方体鬲、扁体鬲这三类陶鬲形态确实无法从最大腹径、通高、档高、实足高这四个变量其中一个指标单独呈现出来的，而且其中又有物比例尺和明器样本，单位不统一。

基于常识和考古报告的描述，完全可以用最大腹径/通高比值作为划分三类陶鬲形态的指标依据，最大腹径/通高比 <1 ，为瘦体鬲；最大腹径/通高比 ≈ 1 ，为方体鬲，最大腹径/通高比 >1 ，为扁体鬲。

(五) 选择合适的统计方法进行统计分析与验证

由于变量数据类型关系为定序变量（时期）——一定距变量（陶鬲形态），应用单因素方差分析进行检验。若两种假说均成立，可以通过统计学相关系数、显著性水平比较，判定哪一种假说更有解释力。略懂统计学常识的人都知道，在某一分类下，样本越多，越容易呈现显著，所以二期说在显著性水平上是有先天优势的。但是相关系数 η 值则与显著性水平无关。

结果如下：

1.对三期说的统计检验：

按照单因素方差分析，首先应进行方差齐性检验，显著性小于 0.01，结果显著，方差不齐，未达到方差分析使用条件，无法使用参数检验，退而求其次，使用非参数检验方法——Kruskal-Wallis 检验。结果如下：

表 5 Kruskal-Wallis 秩检验结果

	三期说	N	秩均值
腹径通高比	一期	49	39.31
	二期	41	79.00
	三期	57	100.23
	总数	147	

（卡方=54.746，渐近显著性<0.01，结果显著，具有样本推论总体的可能）

非参数结果显著，意味着在三期说预测下，不同时期陶鬲形态有统计学意义上的差异；还是给各位读者展现一个简明的三期说与腹径通高比分布图。

2: 对二期说的统计检验

方差齐性检验结果为 0.27，结果不显著，达到使用方差分析的前提条件，方差分析结果如下，这时呈现的统计图才是有意义的。

表 6 二期说与腹径通高比方差分析结果表

腹径通高比					
	平方和	df	均方	F	显著性
组间	1.491	1	1.491	106.009	.000
组内	2.040	145	.014		
总数	3.531	146			

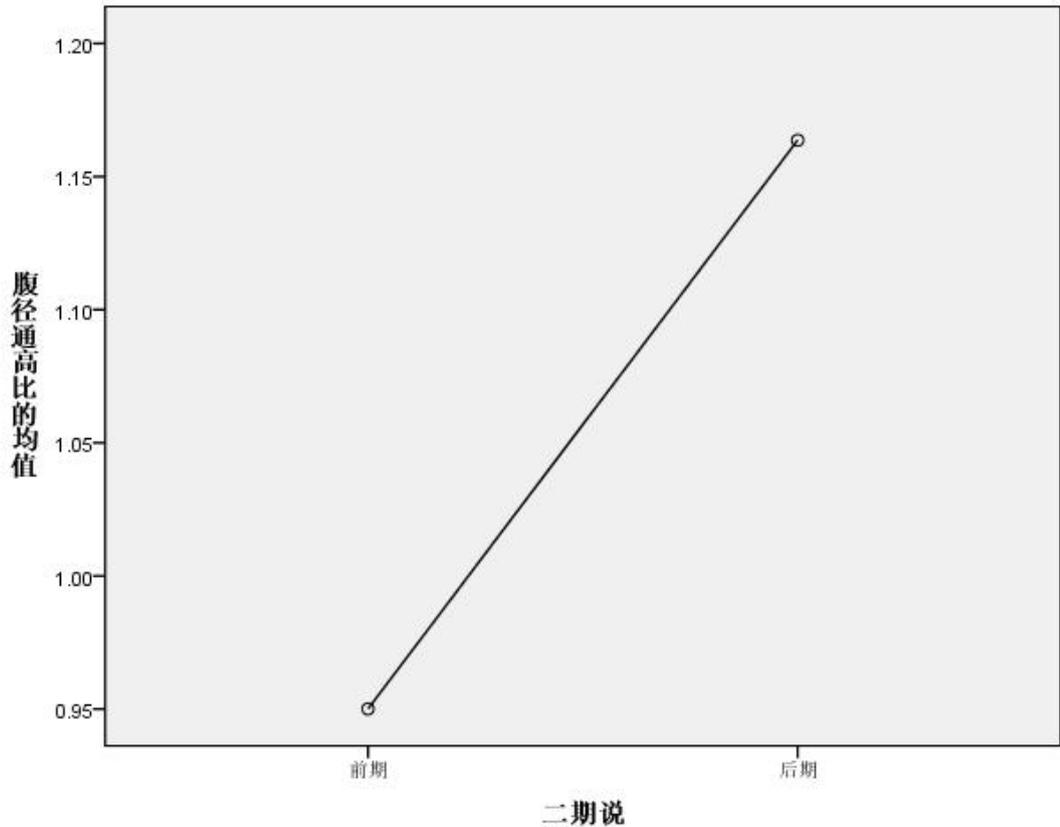


图 2 二期说与腹径通高比分布图

商前期陶鬲腹径通高比均值在 0.95 左右，后期在 1.15 左右，商代陶鬲确实呈现出由瘦高向扁体形态方向变化的趋势，方差分析 F 值小于 0.05，相关系数 Eta 为 0.65，意为已知陶鬲所属时期，推断陶鬲形态的预测准确率有 65%。

综上，虽然三期说通过了非参数检验（K-W 检验），但是参数检验（方差分析）是比非参数检验更具说服力的。

从统计角度而言，考古对于商代陶鬲所呈现出：瘦体鬲→方体鬲→扁体鬲的定性认识，并不准确，商代陶鬲确实呈现出由瘦高向扁体形态方向变化的趋势，但是并不能从统计学上判定存在方体鬲这一理想的中间形态。

三、结论与建议

本文首先对《统计分析在商代陶鬲研究中的应用》一文进行了批判，指出其中统计方法的运用错误，其背后则是考古对统计思想认识上的不足和错误；其后使用该文同样的数据，用陶鬲形态，分别对商文化二期说、三期说进行验证，发现二期说比三期说在相关系数、显著水平、以及参数检验指标上对于预测陶鬲形态更有解释力。

从统计角度而言，考古对于商代陶鬲所呈现出：瘦体鬲→方体鬲→扁体鬲的定性认识，并不准确，商代陶鬲确实呈现出由瘦高向扁体形态方向变化的趋势，但是并不能从统计学上判定存在方体鬲这一理想的中间形态，而仅存在瘦体与扁体两个理想形态。

（一）定量统计思维能够让考古学意识到什么

1. 样本和总体的思维

考古发现的遗存只是样本，从统计角度而言，需要有能否推断到总体进行显著性检验的过程。考古所发现的样本确实很难保证达到统计推论随机抽样的要求，但是样本还是需要检验。

2. 统计分析结果不是补充证据

而不是观察归纳整理结果一股脑的呈现出来，统计分析结果仅仅沦为补充分析证据。结合本研究翻译为通俗易懂的语言就是考古学家运用考古类型学，将商代陶鬲进行了分期，分期仅仅是基于样本提出的假设，假设是否具有统计推断的性质，解释力是多强？推断准确率是多少？可见类型学的分期结果从统计学来看还上升不到观点层面。

2. 从统计思想看考古学分期

本文在进行统计分析时，需要结合变量类型选择合适的统计分析方法，遇到了这样一个问题，考古学对商文化的三期或二期的划分，从统计学角度而言，分期是一个定类变量还是一个定序变量？即不同时期的划分是仅仅有 a 非 b 的区别还是有 $c > b > a$ 的大小区别，分期有时间先后所以是定序变量，那么能否推断商后期的文化比商前期文化繁荣？进而推断扁圆形陶鬲比瘦高型陶鬲所代表的文化期更先进呢？也许可以通过实验证明扁圆形陶鬲比瘦高型陶鬲煮饭更有效率，但这是否就意味着文明有所进步呢？考古学家在使用考古类型学分类的初衷多仅仅是区别，而非定序意义上的不同文明阶段有可比较性的差异，所以在上述研究中分期变量按照定类变量处理，这也就影响到了相关系数的选择。

统计分析的核心有两个，一是变量关系问题；二是统计推论问题；考古类型学仅仅解决了某一变量内部类别划分问题，但这仅仅是非常初步的基础性分析，仅为分类或者仅止步于分类的划分，还算不上真正的研究。

（二）对定量考古规范化建议

1. 系统性学习社科研究方法论为前提去运用统计

当下有很多统计软件教材，但也仅仅是统计软件的操作与简单的解释，没有深入介绍这些统计方法背后的思想和适用预设和前提^[6]；我国考古学界也亦有相关国外考古界翻译教材的引进^[7]、以及学科内定量方法的教授，但这均是二手知识的学习，对于大多数不是理科出身的考古研究者而言，从概率论、线性代数、开始学习并不现实。更为可行的是在社会学、心理学、经济学这些统计学运用较为成熟的社会科学下开展相关定量课程学习，考古学本身不具有完整学科知识体系^[8]，但是运用社科研究方法论与研究方法对考古资料开展研究是最好的选择，定量统计仅仅是社科研究法法之一、它背后亦有相应统计思想支撑。

按照顺序，最基本需要掌握如下知识：社科研究方法和方法论——概率论与数理统计——社会统计学（社科通用）——软件操作相关知识。最好能边学，边有研究进行练手，系统地运用在自己的研究里，否则仅仅沦为纸上谈兵。如本文所展示的那样，统计学最讲究的就是针对不同变量数据类型（如定类——定距），选择合适的统计方法。

如果没有上述学习能力，则可与懂得统计的研究者合作合作。切不可自己没有定量统计思想与判别力就单纯在形式上模仿与复制定量研究，单纯学习统计软件，跑数据并不困难，但是没有统计思想的统计分析只会让定量分析沦为结论的辅助证据呈现工具。如《统计分析在商代陶鬲研究中的应用》的第一段、第二段通过简略提及已有的考古学定量研究，进而确立该文继续进行定量考古研究的合法性，而没有进行符合学术论文规范要求^[9]进行批判性文献综述。

2. 考古学运用定量研究要基于统计思想

正文中第一部分已经提到了统计推断（即社科中的高级统计）是包含假设检验思想的一一样本结论是否能推断到总体结论？考古学文化类型学的分期与分类观点在统计上是否能够得到证实或证伪？

很多研究运用了统计分析,最后仅仅是为了证实或拟合了既有的观点,甚至可以改变操作化关注点,有意拟合结论,而不展示不利于既有结论与分析路径的定量结果,这完全是让统计成了研究的装饰。

这其实与我国考古学研究多停留在资料收集、整理的低端研究现状有关,在考古报告撰写阶段,描述统计确实能够简化资料,让资料更好地呈现在读者面前,如本文研究路径二——聚类分析:如果是考古资料分析阶段,则需要有科学方法论——假设检验的思路,把统计视为检验假设的工具,而非拟合结论的包装,把提出的观点与认识视为还需统计检验的假设。

总之,考古学是通过物质遗存理解古人文化与生活的学科,对物质遗存的任何技术性、跨学科的分析均是以该指向为终极目标的。商代陶鬲的不同形态能说明商文明什么问题?能反映多少商人文化生活状态?还是说考古类型学仅仅只能停留在类型学呢?这是仅靠统计分析无法回答的。

参考文献:

- [1] 叶舒然.定量考古学研究述略——兼谈在青铜器分期研究中的应用[J].学术界,2016(8):239-245.
- [2] 陈铁梅.定量考古学[M].北京:北京大学出版社,2005.
- [3] 刘歆益,朱剑,严艳.统计分析在商代陶鬲研究中的应用[J].南方文物,2007(3):67-72.
- [4] [美]巴比.社会研究方法(第11版)[M].北京:华夏出版社,2009.
- [5] 孙庆伟.鼎宅禹迹[M].上海:生活·读书·新知三联书店,2018.
- [6] 陈铁梅.定量考古学[M].北京:北京大学出版社,2005.
- [7] [美]周南.给考古学家的统计学[M].北京:中国社会科学出版社,2021.
- [8] 宋江宁.无角度、不学术——对中国考古学学科定位和发展方向的思考[EB/OL].中国考古, http://kaogu.cssn.cn/zwb/xsdt/xsdt_3347/xsdt_3348/202006/t20200603_5138430.shtml

作者简介:陈锦航,男,纯社会学学科背景,2017年获北京大学社会学博士学位,原法门寺博物馆副研究员。