

# 数学方法在考古类型学研究中的实践与思考

滕铭予

(吉林大学边疆考古研究中心, 吉林长春 130012)

**中文摘要:** 在考古学研究中加强定量研究的思想已成为当前众多考古学者的共识。定量研究的前提是引进数学方法。考古类型学研究的基础工作就是分类, 分型是分类, 分式亦是分类, 只不过分类的原则和目的不同而已。分类的最根本的目的是搞清研究对象之间的各种关系, 其相似、相异的程度。而数学就是研究事物之间关系的科学, 所以分类本身即是一种数学行为。

本文利用山西省侯马市乔村墓地 1963 年至 1993 年间的发掘资料, 运用多元统计分析方法对墓葬中随葬的陶盆、陶釜、喇叭口斜领罐和小口有领罐进行类型学分析, 以探讨多元统计分析方法应用于考古类型学研究中的可行性。

**关键字:** 数学方法; 考古类型学; 乔村墓地

**文献标识码:** A

—

在考古学研究中加强定量研究的思想已成为当前众多考古学者的共识。定量研究的前提是引进数学方法。而考古类型学研究的基础工作究其大要是对遗存进行分类, 分型是分类, 分式也是分类, 只是不同层次的分类具有不同的分类原则和标准。而分类的目的则是搞清遗存间的各种关系。数学作为一种方法论科学, 其本质就是研究事物之间的各种关系。从这个意义上讲, 类型学研究的本身亦是一种数学行为。

在中国的考古类型学研究实践中, 已有很多学者认识到了加强定量研究的重要性<sup>1</sup>, 并且自觉或不自觉地在研究中运用了数学方法。如高明先生对东周时期的铜壶进行的类型学研究<sup>2</sup>, 林澧先生对东北系铜剑进行的类型学研究等<sup>3</sup>。这些研究由于其采用的数学方法以计算器物某些部位的比例或数值为主, 并按一定的数值的大小进行排序, 所以需要计算的数据量不是很大, 通过手工操作即可以完成。只是这种方法仅适用于对可以表示遗存特征的若干项数值中的某一项进行计算, 若要对遗存的多项特征进行综合分析, 就必须运用数理统计中的多元统计分析<sup>4</sup>。而多元统计分析在计算过程中有大量的数据需要进行计算, 其计算量之大, 不要说由手工操作难以完成, 在个人电脑不甚发达的过去, 甚至需要在大型计算机上方能进行。不过由于近年来计算机技术的高速发展, 现在这种计算在个人电脑上亦能很快完成。由于多元统计分析在社会科学领域的应用非常广泛, 现已开发出供社会科学研究使用的多种计算机用多元统计分析软件包<sup>5</sup>, 可以说在计算机技术应用如此普及的今天, 使用多元统计分析对考古遗存进行定量研究已经具备了较充分的条件。

已有的中国考古学类型学研究中使用多元统计分析进行定量研究的实例, 严格地讲只有一例, 即陈铁梅和何弩所做的用主成份分析方法对夏商时期的陶豆进行的分类<sup>6</sup>。但由于其所做的陶豆数量过少, 只有 13 件, 而且这 13 件陶豆在外部形态上又存在明显的差别, 加之使用的主成份分析方法, 其主要的功能在于归纳和描述不同的分类模式, 所以其研究的意义与其说是在于对陶豆所进行的分类研究, 倒不如说是因为用数学方法和语言正确地表述了用经验观察到的不同陶豆之间的差别, 而向考古界展示了多元统计

分析用于考古类型学研究的可能性。而本文所要讨论的，则是如何将多元统计分析运用于对陶器的类型学研究，实现计算机技术在考古学类型研究领域的应用，使这种可能性成为可行性。

## 二

本文分析的材料是山西省侯马乔村墓地 1963 年到 1995 年间的发掘资料。之所以选择该墓地做为研究对象有以下几个原因：第一，墓葬中的随葬器物多为完整器物，每件器物作为一个单独的个体其损失的信息量很少，对于进行各种器物的类型学研究提供了较好的基础。第二，每一个墓葬作为一个单位，其中所随葬的器物具有很强的共时性，基本可视为同时生产和同时使用的一组器物，对于利用器物组合判断各种类型器物间的关系提供了较好的条件。第三，在一个墓地中的所有墓葬，其间必然具有某种关系。当然一个墓地中各墓葬间的关系可以通过对其空间分布进行研究而得到解释，如有的墓地存在着明显的空白地带将墓地划分为不同的墓群，有的墓地呈明显的分排现象，因此可通过对墓葬的分布、排列进行分析而获得对其相互间关系的认识<sup>7</sup>。墓地中各墓葬之间的关系还有一部分是通过时间序列反映出来，时间序列尽管也可以在空间分布中有所体现，但大多都需要对其内含进行分析，排列出各墓间的相对年代，才有可能对其进行深入的研究。目前已有的对墓地的分期研究，还多是划分出大的期段，对同处一期中的各墓间的相对年代，则难以得出较为明确的认识。笔者此次分析的目的之一亦是在对随葬器物进行的类型学研究基础上，排列出墓地各墓间的相对年代序列，只是由于时间和篇幅的关系，容另文讨论。

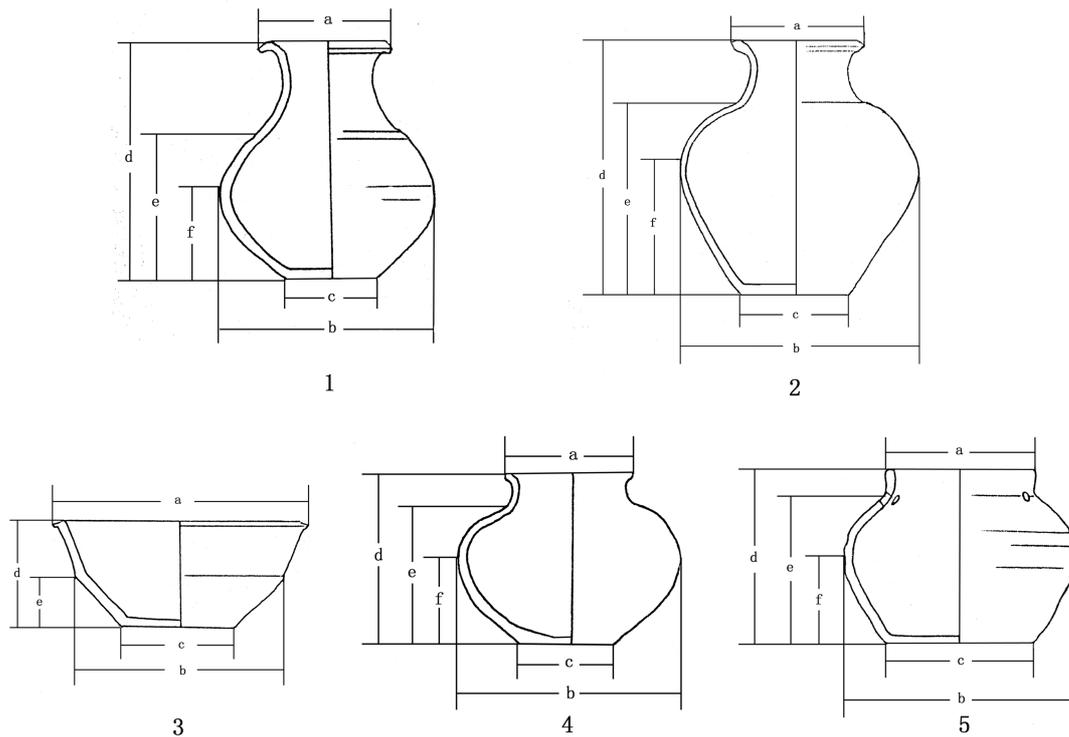
需要说明的是，这批资料目前正在整理，发掘报告尚未发表，笔者所做的工作亦没有受到整理者对这批材料认识的任何影响，即几乎不存在任何先验的认识于其中。也正是由于这种原因，为尊重我国考古界已有的惯例，本文只就此次工作的方法、过程、结论的可信程度等问题进行分析和讨论，而不涉及任何具体资料，如器物测量图，测量数据，墓葬的各种数据等，希望能够得到读者的理解。

利用多元统计分析对器物进行定量的类型学研究，第一步工作是获取原始数据，此次工作所使用的数据主要有二种<sup>8</sup>。

1. 通过对器物的某些特定部位的测量所获得的数值。
2. 不同部位所测得的数值间的比值。

本次分析所使用的绝大部分数据为现场对器物进行实测所得。部分器物因手工制作而形状不规整，如口不圆，高低不平等，考虑到这样的器物并不是制作者有意而为，而是由于在制作过程中出现某些失误造成的，制作者的初衷亦当为形状规整的器物。所以在测量不规整器物时，采用了取最大值和最小值，然后计其平均数的方法。有些数据因测点位置的关系，在实测时不易操作，采用了测器物图的方法获得数据

本次工作共对喇叭口斜领罐、小口有领罐、折腹盆、小卷沿有领罐和直口广肩罐共 5 种器物进行了测量与分析。各种器物的测量部位如图一。



图一 不同种类陶器测量部位示意图

1. 喇叭口斜领罐：a. 口径 b. 最大径 c. 底径 d. 通高 e. 领腹分界点高 f. 最大径高
2. 小口有领罐：a. 口径 b. 最大径 c. 底径 d. 通高 e. 腹高 f. 最大径高
3. 折肩（腹）盆：a. 口径 b. 肩径 c. 底径 d. 通高 e. 肩高
4. 小卷沿有领罐：a. 口径 b. 最大径 c. 底径 d. 通高 e. 腹高 f. 最大径高
5. 直口广肩罐：a. 口径 b. 最大径 c. 底径 d. 通高 e. 腹高 f. 最大径高

第二步工作，对原始测量数据进行分析和整理。从对器物的观察可知，各种器物中不同的个体形体大小虽有不同，但外部形态并没有因形体大小而有明显的差别。考虑到如果直接用测量所得数据进行分析，大小不同的器物会因测量数据出现较大的差值而成为决定分类的主要因素，其结果不仅形态上的差异可能会被忽略，而且在形态上也缺乏可比性<sup>9</sup>。但是若选择各部位测量所得原始数据间能够说明问题的比值进行分析，则可以避免这种偏差。所以在测量原始数据后，还要根据对不同器物的形态进行分析，以确定在分析时所使用的比值。

1、喇叭口斜领罐，154件。喇叭口，较长的斜领，多鼓腹或鼓肩，器高均小于17厘米。通过对其观察发现在形态上的差异主要表现在领的长短和腹部的高矮、腹部最大径的位置、下腹部的收分等部位，据此可以确定下列有意义的比值。

a、长/通高，可以表示领的长短。

b、最大径高/领腹分界点高，可以表示腹部最大径突出的位置，若比值小于0.5为垂腹，位于0.5左右为中腹，大于0.5则形成肩部。

c、最大径/领腹分界点高，可以表示腹部形态是近瘦长还是近方形，亦或是近扁方。

d、口径/最大径，是以最大径作为比较基点来表示口的大小。

e、底径/最大径，可以表示下腹部的收分。

f、最大径/通高，可以表示整体的形状。

2、小口有领罐，102件。小口，小斜领或直领，较高的腹部，最大腹径多偏上形成鼓肩，器高均大于17厘米，多在20—30厘米之间。其形态上的差异主要表现在整体的形状是矮胖亦或是瘦高。据此可以确定下列比值。

a、腹高/通高，可表示腹部占整器高的比例。

b、最大径高/腹高，可表示肩的高度。

c、最大径/腹高，可表示腹部的形状。

d、口径/最大径，以最大径作为比较基点表示口部的大小。

e、底径/最大径，可以表示下腹部的收分。

f、最大径/通高，可以表示整体形状。

3、折肩(腹)盆，22件。腹部有明显折棱形成折肩或折腹。其形态上的差异主要表现在口部或直或敞，折肩点的高矮等部位。可确定下列比值。

a、口径/肩径，可以表示口部形态，比值若大于1，则为敞口，若接近或等于1，则为直口。

b、底径/肩径，可以表示下腹部的收分。

c、肩高/通高，可以表示下腹部的深浅。

d、口径/通高，可以表示相对于口径的盆腹的深浅。

4、小卷沿有领罐，15件。侈口，沿微卷，有小领，领部或有穿，圆腹或鼓腹。器高均小于15厘米。其形态上的差异主要表现在腹的深浅和腹部最大径的位置。

a、领长/通高，可以表示领的长短或腹的深浅。

b、最大径高/腹高，可知腹部形态，是有腹，抑或最大径位置偏上形成肩部。

c、最大径/腹高，可知腹部整体形态是近正方体或为扁方体。

d、口径/最大径，可以表示相对于最大腹径的口的大小。

e、底径/最大径，可以表示下腹部的收分。

5、直口广肩罐，44件。直口接小直领，领部或有穿，平肩或坡肩。其形态的主要差异表现在腹部的深浅及腹部最大径的位置。

a、腹高/通高，可以表示领的长短或腹的深浅。

b、最大径高/腹高，可以表示是鼓腹抑或是有肩。

c、最大径/腹高，可以表示腹部的整体形状。

d、底径/最大径，可以表示下腹部的收分。

e、口径/最大径，可以表示相对于最大腹径的口的口的大小。

本次分析将把已确定的上述各种器物的各种比值，作为进行分析的原始数据。

第三步工作，以器物的各种比值为基础，计算器物两两间的相似程度。多元统计分析在计算器物间的相似程度时，对于不同性质的数据可以用不同的系数表示。对于本次分析所使用的定量数据，通常多以距离系数表示器物间的相似程度<sup>10</sup>。其基本思想是将每一件器物看成是一个个体，每一个体的性质是由所观测的多个数值所表现的，这样也可以把每一个体看成是在多维空间中的一个点，而个体间的相似程度则可以通过点与点之间的距离关系来表述，距离越近，距离系数越小，二者的相似程度越高，反之二者的相似程度越低。距离系数有多种，目前运用最多的是欧氏平方距离系数<sup>11</sup>。

第四步工作，运用多元统计分析进行分析。此次所采用的方法是多元统计分析中的 聚类分析法 (cluster analysis)<sup>12</sup>。聚类分析的基本思想是首先将每一个个体看成是一类，在定义个体间的距离和类间距离的基础上，选择距离最小的两类将其合并成一个新类，再按类间距离的定义计算新的类与其他类的距离，再将距离最近的类合并，如此继续，直到所有的个体合并成一类为止。

由于类间距离可以有多种定义，此次所选择的类间距离定义为离差平方和法，即以合并后同类内标本数值的离差平方和最小为每一次合并的原则。离差平方和法被认为是聚类效果比较好的方法，应用得也比较广泛<sup>13</sup>。

### 三

在聚类分析结束后，需要对聚类的结果进行分析。

在进行聚类计算的过程中，程序会提供聚类过程表和聚类枝状图，但是二者只提供了所有个体在聚类过程中依次合并的过程和最终聚类的结果，但并没有提供把参与分析的所有个体分成多少类的最佳选择。所以要确定最终选择分为多少类，还需要对各种分类的结果进行具体分析。笔者大学时代的老师陈雍先生曾经说过，分类的真谛在于首先要将所有的对象分到不能分为止，然后再将其合并到不能合为止。笔者在过去所进行的考古类型学研究中，曾根据这一分类思想做过充分的实践，深深体会到这一原则的可行性。本着这样一种分类思想，笔者在处理最终选择分成多少类时，主要是观察分类结果是否把所有可能分开的个体全部分开，以此为标准确定基础分类的数量。然后对各类器物进行观察，将若干具有相同或相似特点类再合并到一起，分成若干组。这种组就是类型学研究中器物划分式别的基础。

由于在聚类时所使用的表示器物间相似程度的系数是欧氏平方距离系数，欧氏平方距离系数所能表现的是器物两两间的相似程度，而没有一个固定的比较基点，所以以此为根据进行聚类的结果只是提供了若干个其内部个体高度相近的类，聚类结果所提供的分类以及各类的序号，并不能反映类与类之间的关系及顺序。但类型学研究对器物进行分类的目的不仅在于把相近的器物聚合在一起，即分成若干式别，还要考察不同式别的器物在该器物发展链条中所占据的环节及最终描述该器物发展的过程。聚类的结果以及在此基础上进行的分组只是提供了分式的基础，确定各组在该器物发展序列中的位置以及这一序列的头和尾，尤其是在没有层位关系检验的情况下，则需通过对器物的观察和认识并参考其他诸多因素才能进行。

#### 1、喇叭口斜领罐

根据聚类分析的结果确定基础分类为 22 类，综合比较各类的主要特点，并剔除了个别形态特殊者后，将其分为 5 组，各组所包括原有的基础分类及各组共同特点如下：

第 1 组，12，其主要的特点是喇叭口较大，小短领，腹部最大径位置位于中部。

第2组, 18, 21, 22, 喇叭口较大, 中领, 腹部最大径位于腹中部偏上。

第3组, 6, 9, 10, 11, 14, 16, 17, 19, 20, 共同特点是斜领较长, 腹部最大径位于中部以上, 形成明显肩部。

第4组, 3, 4, 8, 小喇叭口, 斜长领, 腹部最大径位于腹中部。

第5组, 1, 2, 5, 小喇叭口, 斜长领, 腹部最大径位于腹中部以下, 形成明显垂腹。

## 2、小口有领罐

根据聚类分析的结果确定基础分类为12类, 综合比较各类的主要特点, 并剔除了个别形态特殊者后, 将其分为5组, 各组所包括原有的基础分类及各组共同特点如下:

第1组, 3, 12, 圆鼓肩, 整体较矮。

第2组, 1, 2, 5, 斜领较长, 圆鼓肩, 整体较1组稍高。

第3组, 4, 斜领较长, 圆肩, 肩部最大径位置较2组下移, 整体较高。

第4组, 6, 7, 8, 9, 11, 领较3组短, 整体较瘦高。

第5组, 10, 短领, 高腹, 肩部坡缓, 整体瘦高。

## 3、折腹盆

根据聚类分析的结果确定基础分类为11类, 综合比较各类的主要特点, 并剔除了个别形态特殊者后, 将其分为5组, 各组所包括原有的基础分类及各组共同特点如下:

第1组, 1, 侈口, 折腹位置偏下。

第2组, 2, 3, 4, 7, 口微侈, 折腹位置居中。

第3组, 5, 直口, 折腹位置居中或稍偏上。

第4组, 6, 8, 9, 直口, 折肩位置偏上。

## 4、小卷沿有领罐

根据聚类分析的结果确定基础分类为6类, 综合比较各类的主要特点, 将其分为3组, 各组所包括原有的基础分类及各组共同特点如下:

第1组, 3, 4, 领较高, 球形腹。

第2组, 1, 2, 腹部最大径位置较1组偏上, 稍显肩部。

第3组, 5, 6, 有肩, 小底, 下腹收分明显。

## 5、直口广肩罐

根据聚类分析的结果确定基础分类为7类, 综合比较各类的主要特点, 并剔除个别形态特殊者, 可将其分为5组, 各组所包括原有的基础分类及各组共同特点如下:

第1组, 6, 肩部突出, 整体为扁方体。

第2组, 3, 7, 器高较第2、3组更矮, 肩部及下腹收分更加明显。

第3组，4，最大腹径位于腹上部形成肩部，余同第2组。

第4组，1，底径较第1小，下腹收分较大，最大腹径位于腹中部偏上，整体近扁方体。

第5组，2，口径与底径均较大，最大腹径位于腹中部，整体呈正方体。

共有40座墓葬出土上述5种器物中的2种以上，根据其组合关系，并结合部分时代特征明显的器物，可以认定以上各种器物的组序分别为其式别，其发展序列依序号顺序由早至晚，器物的发展趋势亦如组序所表述。根据各种器物的式别及组合关系，可将其划分为六期，各期段器物组合情况见表一，该表亦可视为随葬有上述器物的墓葬的分期。

表一 器物分期示意表

分期	喇叭口斜领罐	小口有领罐	折腹盆	小卷沿有领罐	直口广肩罐
? *	1 ? **	1? ***	1? ***	1 ? **	1? **
一	2	1, 2	1, 2	2	2 **
二	3	2, 4	2		3, 4
三	3, 4	2, 3	2	2	
四	4	4			5
五	4, 5	3, 5			4
六	5	3, 4, 5			4

\*：表中所排的第一期是就有组合关系的墓葬而言，有些器物就其形态观察很可能早于排在表中第一期中的器物，因此就整个墓地而言，不排除还存在着于表中第一期的期段。

\*\*：没有发现与该类器物中的第1组存在组合关系的墓葬，但不能排除其第1组器物存在于早于表中第一期墓葬中的可能性，

\*\*\*：就此类器物而言，其第1组器物与第2组器物共出于表中的第一期中，但不能排除第1组器物亦出现于早于表中第一期墓葬中的可能性。

#### 四

笔者在运用多元统计方法对上述各种器物进行类型学研究的过程中，就每一步工作都曾有过多次的试验分析，因分析条件不同，结果亦有差异，只是由于篇幅有限不能将每一次试验的结果都记录于此。但就具体分析过程而言，有以下问题需要说明。

1. 在测量以及输入原始数据时，一定要保证数据的准确性，因为原始数据是整个分析工作的基础，不论出现任何差错，尽管有些错误所涉及的数值很小，都会给分析结果带来难以解释的混乱。

2. 选择作为分析时使用的比值，一定要具有分类学的意义。如果选择了过多的无意义、或意义不大的比值，不仅不会使分析结果更接近于其所存在的分类，反而会给分析结果带来干扰。

3. 多元统计分析中用于个体分类的方法主要为聚类分析，但聚类分析中亦有多种方法可以选择。笔者在最初曾就一批材料用不同的聚类方法进行分析，并比较其分析结果，以离差平方和法所得结果，其类内个体相近程度最为明显，亦最接近于类型学中分类的原则。但这并不意味着在类型学研究中只能使用聚类分析中的离差平方和法进行分类，应根据分析对象的条件不同和分析目的不同，选择不同的分析方法当然最好能同时使用多种方法进行试验分析，比较分析结果后再选择其中最适宜者作为主要使用的方法。

4. 在具体分析的过程中，笔者发现并不是每一次都可以得到令人满意的结果，其原因除了有原始数据的错误、选择了不合适的比值外，有时是由于分析对象中包含着一些特殊形态的器物，从而给分析结果带来较大的干扰。所以在分析过程中，应注意剔除特殊形态的器物，以保证分析结果的合理性。

5. 多元统计分析作为一种数学方法，只是为我们提供了一种用不同的手段来说明研究对象间关系的工具，可以提供若干种分析结果，但却不能对分析结果作出合理的解释与说明，更不能提出任何考古学研究的结论。因此使用多元统计分析进行考古类型学的分析和研究，并不意味着对原有的类型学研究的方法和经验的摒弃，其本质只是对原有的对研究对象进行直观的、定性研究方法的一种补充，尤其是对分析结果的解释还需要结合已有的经验和认识，结合对器物形态的观察来进行。实际上，笔者在对基础分类进行归纳划分为不同的组别，以及分析有共存关系的墓葬，并确认各组器物组别的序列时，都是依靠对器物外部形态的观察等定性研究的方法而得出最后的结论。因此，本文可以视为是在考古类型学研究中运用定量研究和定性研究相结合的产物。不过本文所做工作毕竟还只是一种初步的尝试和探索，由于时间和篇幅关系，亦未能将全部工作过程表现出来，其结论正确与否，亦未接受检验。不过正是由于在类型学研究中，运用了数学方法进行了定量研究，并由此而引进了计算机技术在考古学研究领域中的应用，不仅可以快速、准确和清晰地处理大量数据，而且还可以使考古工作者有更多的时间对本来的考古学问题进行更充分、更深入的思考，也许这正是将数学方法和计算机技术引入到考古学研究领域中的魅力所在。

附记：本文分析所用资料，为山西省考古研究所侯马工作站提供；吉林大学考古学系硕士研究生洪石同学参加了测量原始数据的工作；吉林大学商学院孟庆福老师为统计分析工作提出了有益的建议；美国 FOUNTAIN 公司副董事长王家声先生提供了分析工作所用的 SPSS 软件，在此谨向曾无私地为此项工作提供过帮助的各位同仁、各位朋友表示最诚挚的谢意。

1998年6月第一稿

2003年5月第二稿

## The Practice and Consideration for Mathematics Method in Archaeological Typology Study

**Abstract:** It is common accepted by many archaeologists today that we should strengthen the investigation of ration in

the study of archaeology. The precondition of the investigation of ration is introducing into mathematical method. The basis of typology study in archaeology is classification, discriminating type is classification and discriminating pattern is also classification. But the principle and intention is different. The most essential intention of classification is beating out different relations as well as their degree of similarity and dissimilarity of the study objects. However, mathematics is a subject researching the relation among objects, so that classification itself is a kind of mathematical behavior.

This paper has used the datum from excavation in graveyard at Qiao Village, Houma City, Shanxi Province from 1963 to 1993. Having wielded Multivariate Statistics method analyses the pattern of ceramic pots, kettles, bell-mouthed skew collar jars and small-bore with collar jars in tombs, in order to discuss the feasibility of multi-element statistics analytical method in archaeological typology study.

**Keywords:** Mathematical Method,; Typology in Archaeology; the Graveyard at Qiao Village.

**基金项目:** 国家教委留学回国人员基金资助项目

**作者简介:** 滕铭予(1954-), 女(满族), 山东省定陶县人, 吉林大学边疆考古研究中心教授, 博士, 1992—1993年赴日本西南学院大学研修, 主要从事中国北方考古、战国秦汉考古学研究。

<sup>1</sup> 陈铁梅, 考古学中的定量研究 [J]. 考古与文物, 1993, (6): 94-100.

<sup>2</sup> 高明, 中原地区东周时代青铜器研究(中) [J]. 考古与文物, 1981, (3) 84-103.

<sup>3</sup> 林澹, 中国东北系铜剑初论 [J]. 考古学报 1980 (2): 139-162.

<sup>4</sup> 滕铭予, 多元统计分析方法及其在考古学研究中的应用 [J]. 考古学集刊, 2000 (13): 309-319

<sup>5</sup> 目前国际广泛使用的多元统计分析软件包主要有: SAS(Statistical Analysis System), 为大型统计分析系统, 由美国 SAS 公司开发, SPSS(Statistical Package for the Social Sciences), 为社会科学统计软件包, 由美国 SPSS 公司开发。二者均可以在各种型号的计算机上运行。

<sup>6</sup> 陈铁梅, 何弩等, 计算机技术对河南省二里头二期至人民公园期陶豆分类的尝试[C]. 考古学文化论集(二), 北京: 文物出版社, 1989.340-345.

<sup>7</sup> a: 北京大学历史系考古教研室, 元君庙仰韶墓地 [M]. 北京: 文物出版社, 1983.

b: 山西省考古研究所, 上马墓地 [M]. 北京: 文物出版社, 1994.

<sup>8</sup> 一般情况下, 数据的取值方法有三种, 第一种, 取值为实数, 对于器物来说则是通过对某些部位的测量而获得的数值, 或不同部位的尺寸间的比率。这种数据被称为定量数据。第二种, 取值只有序列的关系而没有明确的量的表示, 如器物的口沿有敛口、直口和侈口等, 这种变量虽然没有明确的数量表示, 但可以用具体数据来表示, 如不同的口沿可用 1, 2, 3 来表示。这种数据被称为定序数据。第三种, 有些数据既没有明确的数量表示, 又没有序列关系, 只是有或无、是或不是的关系, 如陶器戳印文字的有无, 往往用 0 和 1 来表示。这种数据被称为定性数据。本次工作因实地测量的数据比较充分, 所以均采用定量数据进行分析。

<sup>9</sup> 笔者曾经使用盆的原始测量数据进行过试验分析, 其结果是器体的大小对器物的分类起到了决定性的作用, 而器物的形态差异则成为附属于形体大小之下的附属因素。

<sup>10</sup> 在多元统计分析中, 常用的表示个体间相似程度的系数有距离系数, 多用于定量数据, 匹配系数, 多用于定性数据。另外还有可以表示变量间相似程度的相关系数、夹角余弦系数等。

<sup>11</sup> 欧氏平方距离(squared Euclidean distance)定义为:

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2}$$

其中  $m$  为每个个体所观测的变量的数量,  $x_{ik}$  为第  $i$  个体所观测到的第  $k$  个变量值,  $x_{ji}$  为  $j$  个体所观测到的第  $k$  个变量值。常用的距离系数还有绝对值距离、明考夫斯基距离、切比雪夫距离和马氏距离等。

<sup>12</sup> 在多元统计分析中, 根据分析目的不同可以有多种方法, 主要包括以下几类: 1, 预测: 回归分析; 2, 归类: 判别分析; 3, 分类: 聚类分析, 典型相关分析; 4, 归纳: 主成份分析; 5, 说明: 因子分析; 6, 检验: 方差分析等。

<sup>13</sup> 类间距离是指计算每一新类与其他各类间的距离时所采用的方法, 除了本文所使用的离差平方和法外, 还有最短距离法, 其定义为两类间距离等于两类最近的个体间的距离; 最长距离法, 其定义为两类间距离等于两类最远的个体间的距离; 重心法, 或称均值法, 是以每一类中所有个体的均值作为类的重心, 两类间的距离等于两类个体的重心间的距离; 类平均法, 是以两类个体两两间的距离平方和的平均值作为两类间的距离等。但不论使用何种方法, 在进行每一步的合并时, 都是将类间距离最近的两类进行合并。