

産地推定における統計的手法

東 村 武 信*

1. 序

遺物の属性として、その時間座標である製作年代と同時に、その空間座標すなわち製作地ないしは産出地点を確定することは何にもましてたいせつなことである。古代学に関心を持つ自然科学者は、この産地推定に興味を持ち、発光分析、放射化分析、蛍光X線分析などのいろいろの機器分析法でもって遺物の産地を同定しようとする研究を盛んに行ってきた。これらは、元素含有量の分析により産地同定をしようというものであるが、産地推定のためには、元素分析のみならずも他のいろいろな物理的・化学的性質を指標にすることも当然考えられる。¹⁾そして、これらの「多くの」性質を測って、遺物を産地ごとに分類したり、特定の産地に帰属させたりするのが産地推定の手法である。

本稿の議論は、元素分析を例として進めていく。しかし、他の性質を指標にする場合も数学的手法は全く同じであって、本稿でのA元素という語を、性質Aと読みかえて考えていける。たとえば、ESRを用いて色中心を測ったならば、本稿の中のA元素の含有量というのを、A種の色中心の濃度と読みかえればよい。

なお、この統計的手法というのは、既に多くの人が数理統計学で研究を行ってきたものであり、したがって、本稿は筆者の原著論文ではなく、単なる紹介文にすぎない。筆者が、実際に産地分析を行っていく際に、どんな風にして最もたしからしい結論を得たらよいかにつき勉強した結果を紹介したものである。産地分析にたづさわる人々に、この紹介文を利用していただければと思うのが本稿の趣意であって、そのため、用語に数理的な厳密さに欠ける点があるかもしれない。

2. 産地分析における二つの手法

具体的な例をとって話を進めよう。第一には、石器の石材の原産地の判定である。

筆者たちは、サヌカイト石器原料の産出地の判定を研究している。^{2,3)}サヌカイト原石の原産地は西日本に限られていて、十数ヶ所が知られている。この各原産地からサヌカイト原石を採取し、そ

* 京都大学原子炉実験所、大阪府泉南郡熊取町

これらの元素分析を行うと、たとえば、二上山産の原石の元素組成はこれこれ、香川県五色台産の原石の元素組成はこれこれといった値が得られる。そして、次に、発掘された石器の元素分析を行い、得られた値が、どこの原石のに一致するかということを判定するのがこの方法である。

この場合には、発掘された原産地未知の試料が帰属すべき原石群* が与えられている。すなわち、二上山群という一つの基準群があり、未知試料 α が、この群に属するか否かの検定を行うのである。第二の基準群である五色台群に対しても同じく α がこれに属するか否かを検定する。第三、第四と進めていって、結局、未知試料 α が、どの基準群に属するかが定まる。もちろん、どの基準群にも属さない場合もでてくるが、それは統計的手法の問題ではなく、もっと他の原産地（基準群）を発見しなければならない。

この方法 — 同定法と名付けよう — によると、未知試料 α が、基準群Aに帰属できる確率は何パーセント、Bに帰属できる確率は何パーセントという風に、定量的な判定ができる。基準群AとBとがかなり明確に違った性質をもつものならば、試料 α はAに属する確率が90%くらいでありBに属する確率は0.01%以下であるというような値が得られ、この試料はAのものであると言い切れるのである。

第二の方法は、クラスター分析法であって、分類の方法と呼んだがよいだろう。これは、基準群といった特別のものがいない場合である。上の石器の分析でいうならば、原石試料を採取しないで、言いかえれば、原産地がどこであるかということは問題にしないで、発掘された多数の石器を分析して、これとこれとは同種の原材から、また、これとこれとは別種の原材から作られているという風に分類していく方法である。各地から出土した土器の元素分析を行って、同一原産地のものを選び出すような問題に適用できるのであって、この場合、原産地が「同じ」か否かだけを問題にし、具体的に、原産地が地図上の何処であるかは問わない。石器や土器の型式を数値的に分類していく場合などにも利用できるであろう。⁴⁾

3. 同 定 法

前節で述べたように、同定法では、先づ、一つの基準群Aの性質を知っておかなければならない。そこで、二上山へ行って、サヌカイト原石をたくさん拾ってきて分析する。得られた元素含有量は、同じ二上山といっても、原石試料によって多少はバラツキている。

いま、たゞ一種類の元素、たとえば、カリウム(K)だけを測ったとしよう。そして、二上山の原石No.1試料のカリウム含有量は x_{k1} 、原石No.2のカリウムは x_{k2} といった風に得られた。これら

* 群論で言う群ではなく、常識的な意味で、群と呼んでおく。正確には、母集団と言うべきである。

の平均値

$$\bar{x}_k \equiv \frac{1}{n} (x_{k1} + x_{k2} + \cdots + x_{kn}) = \frac{1}{n} \sum_{i=1}^n x_{ki} \quad (1)$$

が二上山群の原石の中のカリウム含有量の目安となる。こゝで n は、測った二上山原石の数であり、20 ~ 30 ケくらいはとらなければならない。二上山群の中のバラツキは、群内分散

$$\begin{aligned} V_k &\equiv \frac{1}{n-1} \{ (x_{k1} - \bar{x}_k)^2 + (x_{k2} - \bar{x}_k)^2 + \cdots + (x_{kn} - \bar{x}_k)^2 \\ &= \frac{1}{n-1} \sum_{i=1}^n (x_{ki} - \bar{x}_k)^2 \end{aligned} \quad (2)$$

または標準偏差

$$\sigma_k = \sqrt{V_k} \quad (3)$$

で表される。

二上山群の試料のカリウム値が、平均値のまわりにどんな分布をしているかは判らない。この点が問題として残るのであるが、かりに、ガウス分布をしていると仮定しておく*。そうすると、発掘試料 α のカリウム量 x_{k0} の値から、この試料が二上山群に帰属できる確率を求めることができる。すなわち、 $t \equiv (x_{k0} - \bar{x}_k) / \sigma_k$ において、 t 分布の表から、自由度 n のその確率を読取ればよい。

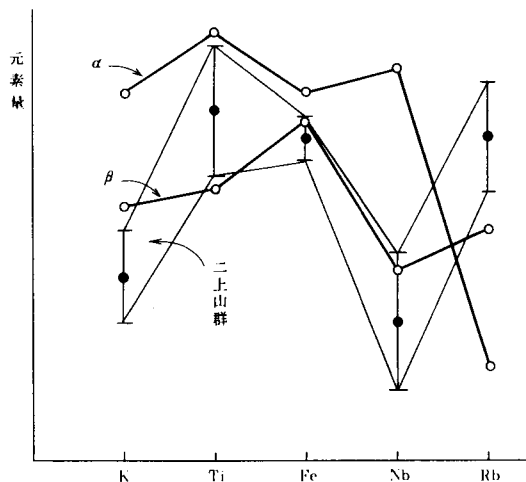
以上は、変量がカリウム含有量というたゞ一種類のみのときのことである。しかし、元素分析は一般には数種類ないしは十数種類の元素量が同時に分析できる。この場合に判定によく使われたのは、次のような、元素含有量のパターンを図に書いて比較する方法であった。

第1図はその模式的な図である。横軸には分析した元素の名前、縦軸にはその元素量をとっておく。二上山群の原石試料の、それぞれの元素の含有量の平均値を図にとって結び、それに、各標準偏差だけの巾をもたして、帯状のグラフを画しておく。発掘試料 α の元素量をこの図の上にプロットして、 α のパターンが帯に一致するかどうかを判定して、 α の帰属をしらべる。

この方法では、第1図の α のような発掘試料ならば、明らかに二上山群原石とは異なると言えるであろう。しかし、 β のような発掘試料ならどうであろうか。 β では、K と Ti と Rb との量が帯から外に出ている。しかしそのずれはわずかである。標準偏差の意味からすれば、たとえ二上山の原石を1ケとってきてもKならKの値が、帯から外へ出る確率が32%ほど存在する。だから、こ

* 85 ページで更に検討する。

のていどわずかに帯の外へ出ていても、二上山群に属しないとは言えない。しかし、一種類の元素だけが外れているのではなく、三種類も外れているのだから、やはり二上山群のものとは違うのではなからうか。結局、このパターン法では、判断できない。パターン法は、操作が非常に簡単で有用ではあるが、差違の大きい試料の場合に使えるものであって、差違の小さい場合には区別できない。そして注意しなければならないのは、小さいながらも差違



第1図 元素含有量パターン

が存在する試料をば、差違がないと誤審してしまい、本当は二上山原石ではないのに、二上山原石を使っていると誤った判断をしてしまうおそれがあることである。

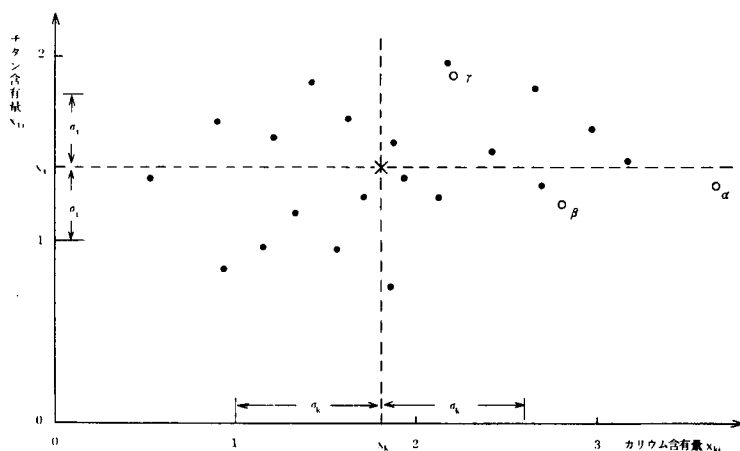
このようにパターンの比較では定量的な判断が下せない。パターン法は、 α や β の試料が二上山群に帰属できる確率が何パーセントであるのかという定量的な結論を導き出せる方法でない。自然科学的方法を考古学へ持込むことの意義は、定量的客観的な結論を導出することと、得られたデータのすべてを最高に有効に使うことである。この二つの立場からすれば、パターン法は不完全な分析であると言える。完全な分析をするには、多変量統計の手法が必要になる。

多変量の統計的解析を進める際、多くの種類の変量を何とかして1ケの変量にまとめなおして、その変量の分布で話が片付くようにもっていくというのが一つの指針である。そうしてまた、多変量の場合には変量間の相関の存在を考慮することが避けられない。

はじめに、相関がない場合、すなわち、変量は互いに独立に分布している場合を考える。図で表すために、変量は2種類の場合をとろう。サタカイトの分析なら、カリウムとチタンとを分析した場合である。二上山の原石を分析して、第2図が得られたとする。

第2図では、各点がそれぞれ、ひとつひとつの原石の分析値であり、平均値の座標線が示してある。チタン量は比較的バラツキが少ないのに対し、カリウム量は大きくバラツキている。実際、カリウムの標準偏差が0.8であるのに対して、チタンの標準偏差は0.4であることが図に示されている。

発掘した試料が、第2図の α であったとする。 α のチタン量は、二上山群の平均値にかなり近い。



第2図 2元素含有量の分布—相関がない場合

もしもチタン量しか測定していなかったならば、二上山原石を使ったものであると判断したであろう。ところが、カリウム量が平均値より大きくずれているから、とうてい二上山原石から作ったものではあり得ない。このようなことがあるからこそ、なるべく多くの種類の元素を分析した方が誤判断の危険が少なくなるのである。

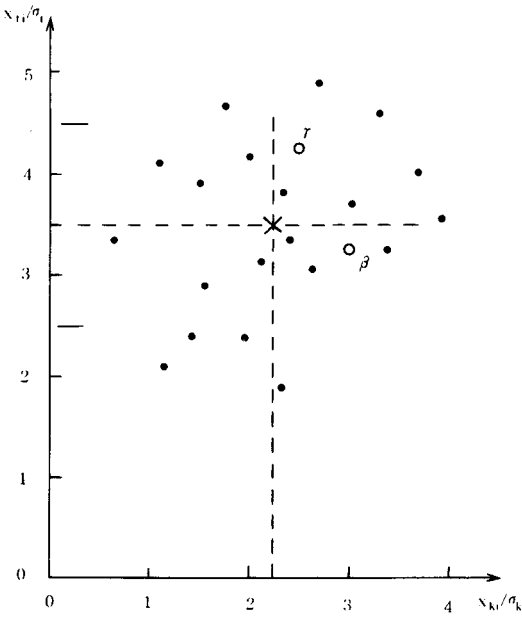
ところで、チタン量の標準偏差は、カリウム量の標準偏差の $\frac{1}{2}$ である。だから、発掘試料を測った場合、カリウム量は平均値 $\bar{x}_k$ からすこしくらい外れていてもよいが、チタン量は、平均値 $\bar{x}_t$ にかなり近くなければ二上山原石とは言えないであろう。図の β 試料と γ 試料とは、原石の分布の中心(X印)から離れている距離は異なるけれども、両方とも同じくらいに、二上山原石からのものであると言うことができる。

だから、試料点の分布を表す図を、縦軸も横軸もそれぞれの標準偏差を尺度として測って描けば、話はもっと簡単になる。すなわち、カリウム量は x_{ki} でプロットせずに x_{ki}/σ_k の値でプロットする。チタン量も同様に x_{ti}/σ_t でプロットする。すると標準偏差は両軸とも1になり、図は、第3図のようになる。点の分布は、円形になるので、試料が二上山群に属するか否かの判定は、「分布の中心からの距離」というたゞ1ケの変数によって行えるであろう。第2図の β 試料と γ 試料とが、第3図では中心から等距離のところに位置している。

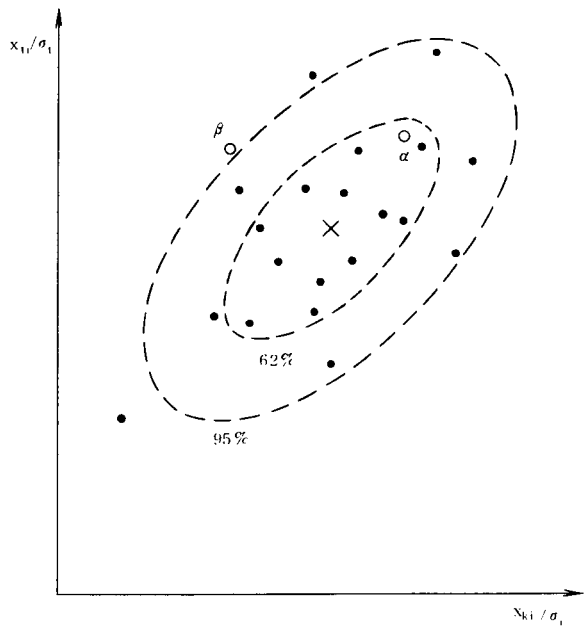
次に、カリウム量とチタン量とに相関がある場合を考えてみる。すなわち、分布が第3図のようではなく第4図のようにになっている場合である。各元素それぞれに、標準偏差の値は前と変っていないが、点の分布は、右上りになっている。これは、カリウムの含有量の大きい試料はチタン含有

量も大きく、カリウム含有量が小さい試料はチタン含有量も小さいという傾向が存在していることを示している。

こういった元素量間の相関は、いろんな天然物の中でいろんな種類の元素の間にひろく存在する。相関が非常に強くて、たとえば、どの試料においてもチタン量がカリウム量のきっちりと0.5倍になっているものならば、チタンは測定しなくても同じである。チタンを測定しても、新しい情報は何も得られない。相関が全くなくて2元素が全く独立に試料に含有されている場合には、チタン測定によって情報量は倍加される。しかし、たいていの場合はこの二つの中間であるので、統計解析は必ず相関を考慮した方法をとらなければならない。



第3図 2元素含有量の分布—相関がない場合



第4図 2元素含有量の分布—相関がある場合

相関が第4図のようになっていて、発掘試料 α と β とが、図の位置に来たとする。分布の中心からの距離は、 α も β も等しい。しかし、図で見るかぎり、 α は二上山群の中に入っているが β は明らかにずれている。だから、「分布中心からの距離」は、判定に適した変量にはなっていない。

第4図の分布は、右上りの楕円形をしている。そこで、適当な楕円を描き、二上山原石試料のうち62%はそこにある楕円、95%はそこにある楕円といった風に描いてみる。すると先の発掘試料 α は、62%楕円領域の中に入っているが、 β は95%領域にすら入っていない。したがって、このような楕円に対応する変量を見つけると、1変量での判定に帰することができる。

右上り楕円に対応する変量をとるには、試料点たとえば α が、分布の中心より右上あるいは左下ならば、中心からの距離をちぎめる方向に補正し、試料点 β が分布の中心より右下または左上にあるときには、中心からの距離を伸ばす方向に補正するような式で表される変量をとってくればよいであろう。それは、数理統計学で言う「マハラノビス距離 (Mahalanobis' generalized distance)」であり、通常、Dをもって記される。この距離は、その2乗が次の式で定義される。

$$D^2 = (x_{k0} - \bar{x}_k)^2 S^{kk} + 2(x_{k0} - \bar{x}_k)(x_{t0} - \bar{x}_t) S^{kt} + (x_{t0} - \bar{x}_t)^2 S^{tt} \quad (4)$$

添字0がついているxは、発掘試料のもの、 $\bar{x}$ は二上山群原石試料の平均値、添字k,tはそれぞれカリウム、チタンを意味している。 S^{kk} などは、二上山群原石試料のデーターの間での分散共分散行列Sの逆行列のkk成分などである。

もし、第2項を省略するならば、Dは、第4図での発掘試料点と中心との距離になっている。第4図では、縦軸も横軸も、それぞれ、標準偏差を単位にとってあるから、 $S^{kk} = S^{tt}$ となるからである。第2項がその補正の役割をするもので、第4図のように右上り楕円のときには(正の相関) S^{kt} は負になる。そのため、第2項全体は、中心から右上または左下の試料点では負、右下または左上の試料点では正になって、楕円の長軸方向の分布が縮み、短軸方向は伸びて、円になってくる。したがって、マハラノビス距離を計算するということは、相関の結果楕円状の分布になっているものを押しまらめて円状の分布にし、そのときの中心からの距離を計算することに相当する。

マハラノビス距離が求まれば、その発掘試料が二上山群に属する確率がわかる。その手続きは本節の最後に述べる。

以上は、2種類の元素の分析結果を解析するときの概念的な説明であるが、もっと多くの種類の元素を分析した場合も、そのまゝの形で理論を拡張することができる。手続きだけを、次に書いておく。³⁾

母群(二上山群)から原石試料をnヶ採集し、それぞれp種類の元素を分析した。いっぽう、発

掘資料 α についても、同じく p 種類の元素を分析した。 $n > p$ であるとする。

i 番目の原石試料の中の、第 k 種および第 ℓ 種の元素含有量を x_{ki} , $x_{\ell i}$ とする。 原石試料中の k 元素含有量の平均値 $\bar{x}_k$, k 元素量の分散 S_{kk} , k 元素量と ℓ 元素量との共分散 $S_{k\ell}$ は次のように与えられる。

$$\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ki} \quad (5)$$

$$S_{kk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ki} - \bar{x}_k)^2 \quad (6)$$

$$S_{k\ell} = \frac{1}{n-1} \sum_{i=1}^n (x_{ki} - \bar{x}_k)(x_{\ell i} - \bar{x}_\ell) \quad (7)$$

分散および共分散の値を、 k ($k=1, 2, \dots, p$) を縦に、 ℓ ($\ell=1, 2, \dots, p$) を横に並べて、分散共分散行列 S を次のように作る。

$$S = \begin{vmatrix} S_{11} & S_{12} & \dots & S_{1p} \\ S_{21} & S_{22} & & S_{2p} \\ \vdots & & & \\ S_{p1} & S_{p2} & & S_{pp} \end{vmatrix} \quad (8)$$

S の逆行列 S' を求める。 S' を求める式は

$$S \cdot S' = 1 \quad (9)$$

である。

$$S' \text{ を } S' = \begin{vmatrix} S^{11} & S^{12} & \dots & S^{1p} \\ S^{21} & S^{22} & & S^{2p} \\ \vdots & & & \\ S^{p1} & S^{p2} & \dots & S^{pp} \end{vmatrix} \quad (10)$$

と書く。

発掘試料 α の中の k 元素量、 ℓ 元素量を x_{k0} , $x_{\ell 0}$ とする。この発掘試料と、二上山群原石試料の分布の中心との間のマハラノビス距離 D_0 の 2 乗は、

$$D_0^2 = \sum_{k=1}^p \sum_{\ell=1}^p (x_{k0} - \bar{x}_k)(x_{\ell 0} - \bar{x}_\ell) S^{k\ell} \quad (11)$$

で与えられる。

$$T^2 = \frac{n}{n+1} D^2 \quad (12)$$

とおき、さらに

$$F = \frac{n-p}{p(n-1)} T^2 \quad (13)$$

とおくと、Fは、自由度(p, n-p)の分散比分布になっていて、Fの表をひけばこの試料が母群に帰属する確率がわかる。こゝで、もし $n < p$ なら自由度が負になってしまう。だから、この解析法は、 $n > p$ すなわち、基準群から採取した試料の数が、分析元素の種類よりも大きいことが条件である。

なお、nが十分に大きいときは、 T^2 は自由度pの χ^2 分布に従うようになり、 χ^2 の表で確率を判断できる。式(11)以下を、Hottelingの T^2 検定とよんでいる。

この節を通じて、元素含有量はガウス分布をしていると仮定していた。しかし実際にはガウス分布をしていないのが普通である。そのとき、元素含有量 x_{ki} をそのままとってこずに、 x_{ki} の対数 $\log x_{ki}$ をばあらためて x_{ki} であるとすれば、この新しい x_{ki} はガウス分布にかなり近いことが多いの例で実証されている。すると、統計解析はこの新しい x_{ki} について本節のやり方で行えばよい。ただし、必ずしも $\log x_{ki}$ がガウス分布をするとは限らないので、⁶⁾ 予めこのことをチェックしておくのが望ましい。

なお、筆者たちのサヌカイト石器の分析では、 x_{ki} としては元素量そのものでなく、2種類の元素の量の比をとっている。これは、非破壊分析のための工夫であって、統計処理法はこの項のまゝでよい。

4. 分類法

この方法はクラスター分析の一つであって、「似たものどうし」をまとめていくという原則に立つものである。この「似ている」度合いを定量化する式はいろいろあるが、どの式がよいかは、先験的には決められない。今まで、考古遺物の産地分析のためにはどの式を用いて成功しているかでもって決めなければならない。こゝでは、米国アルゴンヌ国立研究所の考古化学グループが、放射化分析法で求めた土器の元素分析から産地分析を行って成功した例⁵⁾に沿って説明していく。

土器試料は、ギリシャ、キプロスなどの地点から発掘された前15世紀頃のみケネ式土器である。合計63ケの試料を分析し、K, Feなど17種類の元素の含有量が得られている。

説明の便宜上、かりに、KとNaとの2種類のみの元素が分析されたとする。その元素量を縦軸

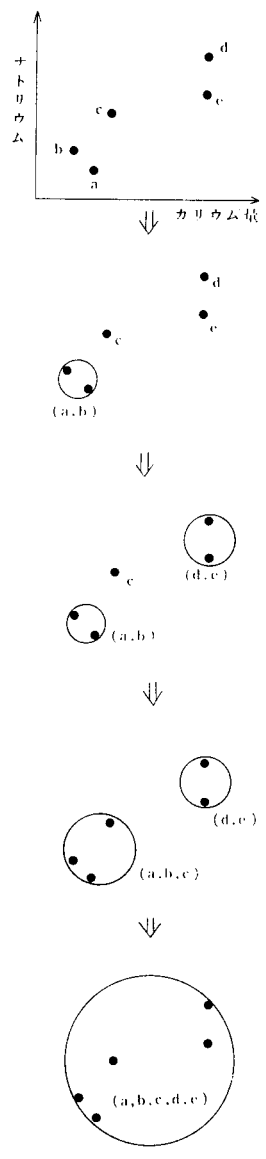
横軸にとって、 $a \sim e$ という5ケの試料につきプロットすると、第5図が得られた。図の上での試料点と試料点との距離に近いほど、両試料は類似している。5ケの試料点を、距離の近いものどうしをまとめ、いくつかのグループ（クラスター）に分類するのである。

このため、先づ、グループの“サイズ”という考えを持込んでおく。たとえば下から2つ目の図のように (a,b,c) と (d,e) という2つのグループに分れるものとするとき、 (a,b,c) グループのサイズとは、その点のバラツキ、すなわち、図に描いた円の大きさである。

始めに帰って、測定データは、 $a \sim e$ の5点であった。この時、 a グループ、 b グループ、……、 e グループという5つのグループであったと考える。どのグループも、ワンマングループである。そしてそのサイズは、どれも、零である。次に、 $a \sim e$ の5点間の距離を、全部計算する。全部で、10ケの距離が出る。このうち、最小のものは、図上で4mmであり、 a と b との組合せである。そこで、この a と b とを一つのグループにまとめる。すると、この結果、5点は、 (a,b) 、 c,d,e という4つのグループになった。このうち、3つはワンマングループ、1つは、2員グループである。最小距離のものを先づ (a,b) とまとめたということは、グループサイズでもって言えば、5点のうち2点を組合せたときにできる2員グループのサイズが最も小さいような2点を選んできたということである。

このように、グループを作ってそのグループのサイズが最も小さいものを選んでくるという原則で順次、一つずつ、点を組入れていく。すなわち、この4グループの中で、どれかをまとめたとき、そのグループのサイズが最も小さくなるのは、今度は d と e との組合せである。その結果、 (a,b) 、 c 、 (d,e) という3グループになった。

その次には、同じようなことを考える。すると、 c を (a,b) と一緒にして、 (a,b,c) というグループ



第5図 分類の順序

プにしたのが最小である。(a,b,c)のサイズは、cを(d,e)と一緒にしたときにできる(c,d,e)グループのサイズや、(a,b)と(d,e)とを合併してできた(a,b,c,d)というグループのサイズよりも小さいからである。これで、(a,b,c)、(d,e)という2つのグループにまとまった。最後に、この2つのグループをまとめ、唯一つのグループ(a,b,c,d,e)にする。

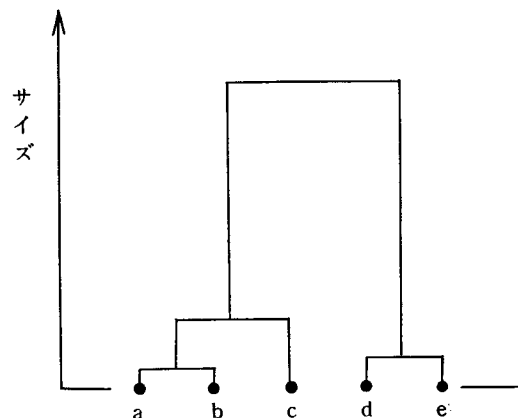
以上がグルーピングの操作であり、1回の操作ごとにグループ数が1つづつ減っていき、全部が一つのグループにまとまるまでつづける。

つぎに、分類樹(系統樹)を作る。それには、最初、全部のa~eを一線に並べる。次に、第1回操作のときにまとまるaとbとを結ぶが、そのときの結合線を、グループ(a,b)のサイズの大きさに相当する高さで水平に引く。(第6図)。第2回操作は、(d,e)を作るものであり、やはり、グループ(d,e)のサイズの高さで、dとeとを結ぶ横線をひく。第3回操作は、(a,b)とcとを一緒にして(a,b,c)を作るものであったから、(a,b,c)のサイズの高さで、(a,b)とcとを結ぶ線をひく。第4回操作は、(a,b,c,d,e)のサイズの高さで、(a,b,c)と(d,e)とを結ばばよい。

でき上がった分類樹を見て、分類するのであるが、いろんな分類で考えられる。すなわち、(a,b)、c、(d,e)と3種に分類するか、(a,b,c)、(d,e)と2種に分類するか、また、(a,b,c,d,e)とみんな同種のものであるか。この他にも考えられる。この判定には、グループのサイズが急に大きくなる所で別種のものとみなすと考えていくのが自然である。したがって、図の例では、(a,b,c)と(d,e)のように2種に分けるのがよい。これで分類が完了した。

この分類操作には、統計の考え方は全く入ってない。上例で、2種に分類するのが「自然である」と言っても、これでなければならぬ理由は何もない。

しかし、いったん分類したものが、妥当な分類であるかのテストはできる。それには、(a,b,c)なるグループを標準群として、dがそれに属するか否かを、dから群(a,b,c)までのマハラノビス汎距離を計算して、前節と同じように判定する。eについてもそれを行う。さらにまた、群(a,b,c)自体についても、その中のaを取出し、aから、残りの(b,c)群までのマハラノビス汎距離を計算して、この距離が十分に小さいことをたしかめる。このような計算を、他の試料についてもすべて行って、たしかめていくのである。



第6図 分類樹

しかし、実際にはこのテストは、完全には行えない。マハラノビス距離による同定テストには、85ページに述べられたように、その群に属する試料数が分析した元素種の数よりも大きくなければならないという条件があるからである。したがって、かなり多数の試料を分類する場合でなければ、この統計的手法による分類の確認はできない。

以上の説明では、直観的な理解に重点をおいたため、正確な表現をとらなかった。グループのサイズを求める式を、最後に示しておく。もっと他のとり方もあるが、下のは、文献5)で採用されているものである。

n 回の試料点で作られたグループのサイズ S_n は

$$S_n = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_{i,j} \quad (14)$$

$$d_{i,j} = \frac{1}{p} \sum_{k=1}^p (x_{ki} - x_{kj})^2 \quad (15)$$

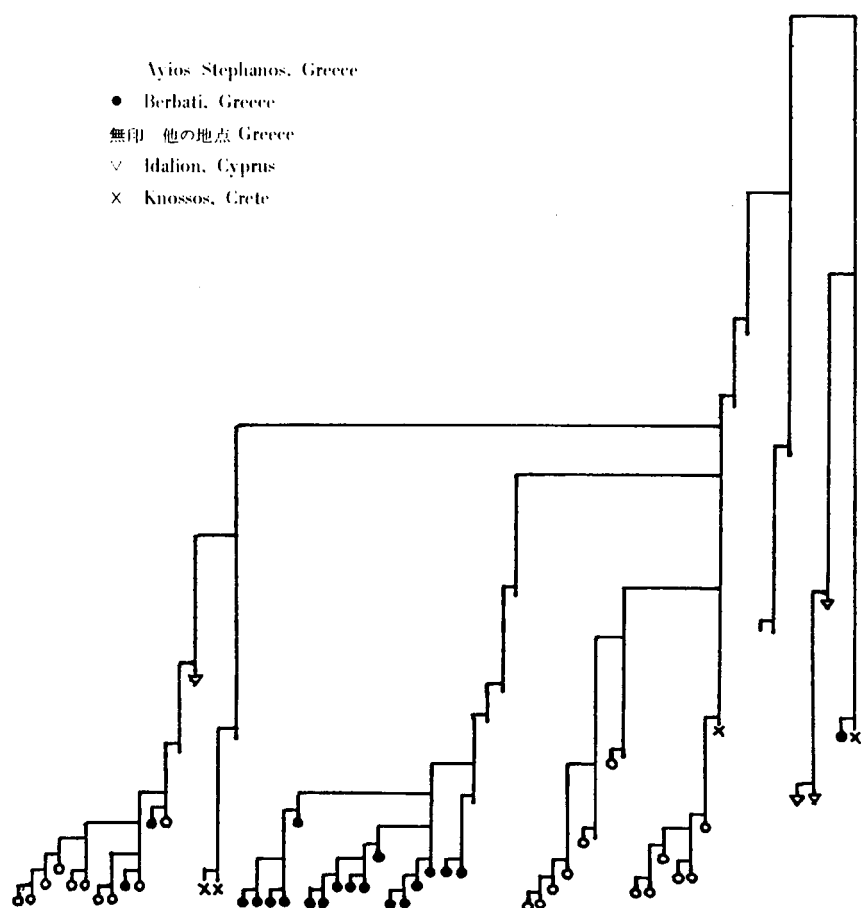
で定義される。 $d_{i,j}$ は i 番目の試料と j 番目の試料との間の平均二乗距離であり、 S_n はちょうど n 回の試料点からこのグループの重心までの距離の2乗和に等しくなっている。

この分類法で行われた土器分類の成功例の分類樹を、第7図に引用しておく。分類判定にたいせつなのは、分類樹の横線の高さであって、試料マークの位置ではないために、試料マークは高さをずらしてある。この結果では、Berbati 出土のものはほとんど全数が同一グループであるが、Ayios Stephanos 出土のものは、3つのグループに分れている。他地点出土のものは、たとえばキプロスで出土したものはこれらのグループとは一致しないから、ギリシャ本土のこの地点からわたってきたものではない。

5. む す び

本稿で述べた2種類の方法を実際に使って分析していく際に、計算に労力がかかる部分は、同定法では逆行列の計算、分類法では、各段階でサイズ最小のものを選ぶ計算である。前者は、 p 種類の元素を分析したときには p 元連立一次方程式を解くことであり、後者は、 n 回の試料のとき $n(n-1)$ 回の S_n 計算を行うことである。計算機のプログラムをいったん組んでおけば、これらは簡単なものであるが、卓上計算機ではすこし厄介である。しかし、といて、計算不可能なほどの問題でもなく、おそらく、数日間の卓上計算で結果が求まるのが普通であろう。

産地分析に限らず、古文化財研究では、分類の数値化など、近い将来にたいせつになってくるものと思われる。いっぽう、物理学や化学の研究者たちは、案外、データーの統計的处理に弱い。この稿で紹介した2方法は、どちらも、完結したものでなく、もっと正確な方法がとり上げられなく



第7図 分類の例

てはならない。たとえば、各測定値の信頼度（または統計的変動）を考慮に取入れた統計的解析法は是非とも採用されなければならないが、筆者の勉強がまだそこまでいっていない。この稿は、筆者の現在の知識の範囲内のものであって、同時に、産地分析などを遂行するには少くともこゝまでは使わなければならないという方法である。

参 考 文 献

- 1) 東村武信, 考古学と自然科学 3 59 (1970)
- 2) 藁科哲男, 考古学と自然科学 5 69 (1972) ; 藁科哲男, 東村武信, 全 6 33 (1973),
全 8 61 (1975)
- 3) T. Higashimura and T. Warashina, *J. Archaeol. Sci.*, 2, 169 (1975)
- 4) F. R. Hodson, *World Archaeology*, 1, 90 (1969)
- 5) A. M. Bieber Jr., D. W. Brooks, G. Harbottle and E. V. Sayre, *Archaeometry* 18,
59 (1976)
- 6) M. de Bruin, P. J. M. Korthoven, A. J. v. d. Steen, J. P. W. Houtman and R. P. W.
Duin, *Archaeometry* 18, 75 (1976)

Statistical Methods in Allocation Problem

Takenobu HIGASHIMURA

Research Reactor Institute, Kyoto University, Kumatori, Sennan, Osaka

In order to find the source location of various implements in archaeological problems, it is very important to analyze data by means of multivariate statistics. This article is aimed to show statistical methods compactly.

The pattern comparing method is very simple but it can not give quantitative conclusions. Hotelling T^2 test using Mahalanobis' distance is useful in many cases. Conclusions obtained by this method are rigid. Clustering technique is also useful to classify implements into some groups although it does not deduce statistical conclusion in the strict sense.