

混合個体群の分析方法と魚類群衆への応用

福 田 嘉 男

Analysis of Intermingling with an Application to Ecological Study of Fish Populations

Yoshio Fukuda

This study is concerned with a method of analysis of intermingling, essential for ecological study of fish populations, by means of characteristics of which frequency distributions are more or less overlapped among the component populations considered. And the present status of research on intermingling of Asian and North American salmon on the high seas is discussed and evaluated from a methodological point of view.

First of all are discussed criteria for selection of characteristics. Precision of the estimated rate of intermingling, as one of them, depends not only on degree of overlap but on symmetry of overlap and the rate of intermingling itself, that is, there is no character best in precision for any situation. More essential criterion is estimability of statistically definite rate of intermingling: the number of classes of the character be not less than the number of the component populations. When this condition is satisfied, the best estimates are provided by method of maximum likelihood, and mechanically by method of scoring. In order statistically to check the model, it is further necessary for the number of classes to be more than the number of the component populations.

Otherwise, when the definite estimation is impossible, very often as it is encountered in ecological study, it is important to evaluate the possibly maximum and minimum rates of intermingling to make a step farther in research. The method of evaluation is presented with application to intermingling of sockeye salmon of Asian and North American origin, of which the present research is considered to be far away from providing any definite estimate of the rate of intermingling.

目 次

I 緒 言	2 頁
II 標識の選択基準 (1)推定混合割合の精度	4 頁
III 標識の選択基準 (2)統計的に確定した混合割合推定の可能性	7 頁
IV 最尤法による確定的な混合割合の推定	9 頁
V 適合度その他の仮設検定と成分個体群に関する 不確かさに由来する推定混合割合の不確定性	11頁

VI 論議	北洋におけるアジア系、北米系ベニザケの混合に関する諸研究の方法論的論議と 推定混合割合の不確定性の分析	13頁
要	約	20頁
文	献	21頁

I 緒 言

ここ数年来実施されてきた日米加3国の共同調査によって、北太平洋およびベーリング海で、夏季に、アジア系、北米系のサケ・マスが、漁業条約締結のとき予想されていた以上に広汎に、混合している事実が明らかにされている。現在、日本は、条約に基づいて、北米系サケ・マスの漁獲を自発的に抑止しており、規制措置として、略、西経175度以東の水域におけるサケ・マス公海漁業は禁止されている。そして、この措置の妥当性の検討、更には、北米系ばかりでなく、アジア系資源の合理的開発のためにも、公海における混合状況を、より定量的に解明することが、不可欠の課題になっている。そのために、共同調査では公海における相当の規模の標識放流とその再捕状況の研究の外に、形態学的諸形質の一次判別函数による分析、鱗の成長特性、ある種の寄生虫、あるいは、血清の抗原抗体反応の有無などを標識とする分析が実施されてきている。この論文は、この課題に関連した方法論的研究がその中核をなしている。

この北洋サケ・マス混合問題は、生物学的な観点から、水産資源研究で種族分析と称されている分野の典型的な問題と見做されよう。水産資源の研究は、漁獲量あるいは漁況変動に見られる諸現象をより本質的に理解するために、それ自身で再生産し世代を重ねるに充分な個体の集り、数量変動の単位としての個体群を究極のところ、研究対象としている。このような単位としての個体群であるか否かを分析し確認する一つの手段として、形態学的形質による調査対象個体群の比較研究が古くから実施されてきているし、研究報告も多数にのぼっている。勿論、形態学的にも、また、分布回遊など生態学的にも、単一の個体群と見做し得るならば、問題はないが、形態学的な異質性を認めた場合、それを指摘するだけでは、水産資源研究の一環としては充分でないことはいうまでもない。同時にまた、たとえ、形態的異質性が認められるとしても、生態学的に、その生活史を通じて隔離している幾つかの個体群については、その形態学的な相違が分類学的関心の対象になり得るとしても、水産資源研究としては問題にはならないことも明らかであろう。水産資源の研究で、所謂種族分析が、課題として重要性を持つのは、そこに、必ずしも均一でない幾つかの個体群の、ある意味での混合が見られるか、予想される場合である。ここで、ある意味での混合というのは、水産資源の研究は、多くの場合、漁獲魚を通じて行なわれ、従って、微細に観察し得るならば、あるいは棲分けを認め得るかも知れないような場合をも、現象的には混合していると思ふべきであるという事情を指している。既往の諸研究の多くは、異質な個体群の存在を指摘するだけで停滞しているように見えるが、北洋のサケ・マスの問題では、異質な個体群の存在は先験的に明らかであり、従って、多くの種族分析が停滞しているところから、研究が始まっているといえよう。勿論、唯単に、個体群の異質性を指摘するだけでなく、たとえ予備的であるにしろ、考察が混合の問題にまでおよんでいる研究もある。方法として見ると、その一つは個々の調査対象群の形質平均値を新たに標識と見做して、群別することである(相川)。形質平均値は、確かに、個体群が定義されて意味をもつから群属性と見做していいであろう。しかしながら、それは、生活史を通じて生物学的内容を持つ個体群が定義された場合であって、個々の調査対象群は、若し混合を予想するならば、変化するもの、もっと任意性をもつものと考えらるべきであろう。生物としてそれ以上細分不可能な個体と同様に、相互に緊密に集団として行動している個体群も勿論、存在していよう。しかし、それは分析の結果として確認されることであり、方法としては、形質平均値を新たに標識と見做して調査対象個体群を分類することは、余りに偏った前提に立つことになるであろう。もう一つは、後に論ずるように、個体の分類に基礎をおく方法である。1957年、「魚類の subpopulation の研究」と題して、所謂種族分析に関する、論文集形式によるシムボジアムが、Marr, J. C. の編輯の下に行なわれている。そこには、現在まで使われ

てきた手段と方法が良く要約されてはいるが、残念なことに、個体群の混合を分析する問題に関して、積極的な展開が見られないのである。

このシムポジウムで、Royce, W. F.^{12a)} が要約している距離函数および一次判別函数による分析の方法は、北洋のサケ・マス、殊にベニザケの混合の分析に、Fukuhara et al⁴⁾⁵⁾¹⁵⁾ も用いているが、それらは、元来、種族間の生物学的近縁性とか、あるいは、新たに発見された化石の分類とか、分類学的な問題を背景にしている。しかし、水産資源研究では、種族というような生物学的に確立された個体群の比較が課題ではないし、また、ある個体を、既知の個体群の何れかに分類することに重点があるわけではない。勿論、混合割合の推定には、混合事象に関与している成分個体群の何れかに、標本の各個体を分類することが、概念的に前提されている。従って、混合割合の推定と個体の分類とは無関係ではないが、自ら相異なる事情もある。この事情は、標識の選択基準に関連して第II章で論ずるが、例えば、個体の分類にたとえ過誤を伴うとしても、もし、相互に、同数ずつ分類を誤るとすれば、混合割合の推定としては差支えないであろう。

さて、ここに、例えば、カタクチイワシの個体群があって、その脊椎骨数の頻度分布が次のようだったとしよう。

脊椎骨数	44	45	46	47
頻度分布	0.05	0.62	0.32	0.01

もし、ある数の脊椎骨をもつ個体の集りを夫々成分個体群と定義すれば、この頻度分布自身、それら成分個体群の混合割合を示している。このように解釈しても、いま定義したような成分個体群に生態学的な意味を見出し難いから何の意味もないのであるが、この例で明らかのように、このような解釈ができるのは、観測標識と分類標識とが一致しているからであり、例えば、異種個体群の間には、そのような標識が存在しているわけである。ここで研究しようとしているのは両者が一致しない場合の問題であるが、そのような標識が存在するかしないかは、成分個体群を、研究の目標として実践的にどのように定義するかにも係っていることを、この例は示している。例えば、年令という標識は、各年令群を成分個体群と定義して分析すべき問題ならば、分類標識とも一致する。しかし、北洋のアジア系、北米系ベニザケでは、年令組成が相違しており、混合の分析に用いられるが、この場合は、年令という観測標識は分類標識と一致していないわけである。

観測標識と分類標識とが一致しないということは、換言すると、観測標識のある実現値をもつ個体が、幾つかの成分個体群に共通して存在することであり、観測標識の頻度分布からいえば、分布が個体群の間で重複しているということである。人為的に標識を付ける場合でも、もし、考える成分個体群の総ての個体に、標識を付けるのでなければ、無標識の個体が、どの成分個体群にも存在し、頻度分布としては重複があることになる。

一般的に、観測標識の頻度分布が X_{ij} ($i = 1, 2, \dots, I; j = 1, 2, \dots, J; \sum_j X_{ij} = 1$) である J 個の成分個体群が、夫々 p_j ($j = 1, 2, \dots, J; \sum_j p_j = 1$) の割合で混合しているものとすると、この混合個体群における観測標識の頻度分布は、

$$Y_i = \sum_j X_{ij} p_j \quad (i = 1, 2, \dots, I)$$

で表される。ここで、 $p_\alpha = 1, p_j = 0 (j \neq \alpha)$ とすると、 $Y_i = X_{i\alpha}$ で、この混合個体群は、 α 番目の成分個体群に外ならない。すなわち、どの成分個体群も、夫々ある特定の混合割合をもつ混合個体群と見做し得る。

この模型は、一見背景の相違している次の問題に見出される。それは、製品検査で、検査規格が例えば定性的で、検査に誤りを伴うときに、正しい仕切不良率を推定する問題である(増井, 1952)。しかしながら、次の2点で、ここで論ずる問題とは事情を異にしている。製品検査では正しく分類されなければならない階級数(分類標識の階級数すなわち成分個体群の数)と、過誤を伴いながら実際に分類される階級数(観測標識の階級数)とが、必ず同じである。しかし、資源研究では、成分個体群の数と、観測標識の階級数と

* 原報告未見。統計数値表に一部引用されている。19頁

は必ずしも一致しない。また、製品検査では、仕切不良率が大きく変化しないように、生産過程は管理されているが、資源研究では、混合割合は時間的にも空間的にも相当に変化すると想定していなければならない。

このような事情から、第三章で、統計的な意味で確定した混合割合を推定し得る条件を論じている。この条件は、第二章で論じている推定量の精度という標識選択基準とは別の、より基本的な標識選択基準であり、従来、看過されていた様に思われる。第四章では、確定した推定が可能である場合、標本の提供する情報を完全に利用する方法として、最尤法による解き方、その実際的な方法として、method of scoring について述べている。

最尤法によって、統計的に確定した混合割合を推定し得る場合、それに付随して模型の適合度、混合割合などパラメーターに関する仮説を検定する問題がある。第五章でその方法と条件について論じている。統計的方法は、普通、統計的に確定した推定が可能の場合に論議が限定されている。しかし、標識自身、あるいは成分個体群に関する不確かさから確定的な推定が不可能になる—これを構造的な不確定ということにするが、このようになる事情は、生態研究では、起こり勝ちであり、調査研究を一步前進させるためには、この構造的な不確定さを明確に分析することが必要であり、重要であろう。第六章では、北洋におけるアジア系、北米系のベニザケ混合問題を素材として、その方法を具体的に述べると共に、以上の所論を総括的に論じている。

II 標識の選択基準 (I) 推定混合割合の精度

混合割合を推定する場合、異種個体群の混合のときにそうであるように、観測標識と分類標識とが一致しているような標識を見出し得るならば、それを用いるのがいいことはいうまでもないことである。そのような標識を見出し得ないならば、どの標識をとっても、成分個体群におけるその頻度分布には、相互に重複があり、その標識を用いると、個体の分類に際して過誤を敢えてしなければならない。Fukuhara et al の用いている Fisher, R. A. の一次判別函数は、幾つかの標識について、2つの個体群の間で、頻度分布の重複がある場合、標本個体を分類する過誤を最小にするような、標識の一次函数として求められている。このような性質の一次判別函数を、新しい標識として採用しても、頻度分布の重複が残ることはいうまでもないであろう。

ある観測標識階級の頻度が、成分個体群 I で X_{i1} 、成分個体群 II で X_{i2} 、 $X_{i1} > X_{i2}$ として、この階級の個体を成分個体群 I に分類するとすれば、任意にとり出した個体が、成分個体群 II に属するにも拘らず、誤って成分個体群 I の個体と分類される確率は X_{i2} ということになる。分布の重複度を $\sum \min(X_{i1}, X_{i2})$ で定義し、常識的に、階級 i の個体を $\max(X_{i1}, X_{i2})$ の成分個体群に分類するとすれば、分類を誤る確率は、分布の重複度に外ならない。混合割合の推定に際して、この分類の過誤ということがどんな意味を持つかを明らかにするために、具体的な例を考察しよう。

人為的に標識をつける場合でも、前述のように、分布に重複があるのが普通であり、任意にとり出した個体を分類しようとするれば、分類を誤る危険をおかさなければならない。勿論、慎重に分類を保留することも考えられるが、例えば、入学試験などでは及落を保留しておくわけにはゆかない。他方、混合割合の推定のためには、標識を付された個体だけを用いること、すなわち、暗黙に前提されている個体の分類という手順に伴う過誤を、全く避けることも可能である。何れにしろ混合割合の推定としては、最終的には、推定量の精度の良否が評価の基準であることはいうまでもないであろう。

いま、成分個体群 I、II および混合個体群における標識の頻度分布が次のようだったとしよう。

	成分個体群 I	成分個体群 II	混合個体群
無標識	X_{11} 0.85	X_{12} 0.75	Y_1
標識 I	X_{21} 0.15	X_{22} 0	Y_2
標識 II	X_{31} 0	X_{32} 0.25	Y_3

そして、比較のために次のような4つの方法を考えよう。

(1) 標識 I, II の階級頻度だけを用いる場合 各成分個体群に異なる標識が付されていて、分類の過誤を全くおかさぬ方法である。このとき、混合割合の不偏推定量は

$$\hat{p}_1 = \frac{1}{X_{21}} \hat{Y}_2; \hat{p}_2 = \frac{1}{X_{32}} \hat{Y}_3$$

もし、 $X_{21} = X_{32}$ ならば、すなわち標識率が等しければ、 $X_{21} = X_{32} = \hat{Y}_2 + \hat{Y}_3$ 。すなわち、全く標識をつけた個体だけで、しかも、成分個体群における標識頻度分布について具体的に何も知るこなしに、推定可能である。公海における標識放流は、同じ標識を付けることで、原理的には標識率を等しくしているわけである。

前述の推定量の分散 $V(\hat{p}_j)$ は、標本の大きさを n として、

$$nV(\hat{p}_1) = \left(\frac{1}{X_{21}}\right)^2 Y_2(1 - Y_2); nV(\hat{p}_2) = \left(\frac{1}{X_{32}}\right)^2 Y_3(1 - Y_3)$$

で与えられる。

(2) 標識 I, II の階級頻度を加え合わせる場合 何れの個体群にも同じ標識を付けた場合である。このとき、同じ標識率ならばいうまでもなく、推定不可能である。 $X_{21} \neq X_{32}$ として、混合割合の不偏推定量は、行列で表わすと、

$$\begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \end{pmatrix} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} + X_{31} & X_{22} + X_{32} \end{pmatrix}^{-1} \begin{pmatrix} \hat{Y}_1 \\ \hat{Y}_2 + \hat{Y}_3 \end{pmatrix}$$

すなわち、 $\hat{p}_1 = \frac{1}{\Delta_2} \{ (X_{22} + X_{32}) \hat{Y}_1 - X_{12} (\hat{Y}_2 + \hat{Y}_3) \}$

$$\hat{p}_2 = \frac{1}{\Delta_2} \{ -(X_{21} + X_{31}) \hat{Y}_1 + X_{11} (\hat{Y}_2 + \hat{Y}_3) \}$$

$$\Delta_2 = X_{11}(X_{22} + X_{32}) - (X_{21} + X_{31})X_{12}$$

そして、その分散は、

$$nV(\hat{p}_1) = nV(\hat{p}_2) = \frac{1}{\Delta_2^2} Y_1(1 - Y_1)$$

で与えられる。

(3) 無標識と標識 I の階級頻度を加え合わせる場合 この場合も次の(4)の場合も、何れか一方の成分個体群にだけ標識をつける場合である。このとき、

$$\begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \end{pmatrix} = \frac{1}{\Delta_3} \begin{pmatrix} X_{32} & -X_{12} - X_{22} \\ -X_{31} & X_{11} + X_{21} \end{pmatrix} \begin{pmatrix} \hat{Y}_1 + \hat{Y}_2 \\ \hat{Y}_3 \end{pmatrix}$$

$$\Delta_3 = (X_{11} + X_{21})X_{32} - X_{31}(X_{12} + X_{22})$$

$$nV(\hat{p}_1) = nV(\hat{p}_2) = \frac{1}{\Delta_3^2} Y_3(1 - Y_3)$$

(4) 無標識と標識 II の階級頻度を加え合わせる場合

$$\begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \end{pmatrix} = \frac{1}{\Delta_4} \begin{pmatrix} X_{22} & -X_{12} - X_{32} \\ -X_{21} & X_{11} + X_{31} \end{pmatrix} \begin{pmatrix} \hat{Y}_1 + \hat{Y}_3 \\ \hat{Y}_2 \end{pmatrix}$$

$$\Delta_4 = (X_{11} + X_{31})X_{22} - X_{21}(X_{12} + X_{32})$$

$$nV(\hat{p}_1) = nV(\hat{p}_2) = \frac{1}{\Delta_4^2} Y_2(1 - Y_2)$$

* この場合抽出変動によって $\hat{p}_1 + \hat{p}_2 = 1$ に見掛上らないことがある。 \hat{p}_1 として $1 - \frac{1}{X_{32}} \hat{Y}_3$ を採用することもできるし、一般に $a_1 + a_2 = 1$ として、 $a_1 \hat{p}_1 + a_2 (1 - \hat{p}_2)$ を成分個体群 I の不偏推定量に選び得るが、精度は元の何れか一方の推定量より良くはならない。

これらの分散に含まれる Y_1, Y_2, Y_3 は任意に変化するわけではなく、次のような混合割合 p_1 ($p_1 + p_2 = 1$) の函数である。

$$Y_1 = X_{12} + (X_{11} - X_{12})p_1$$

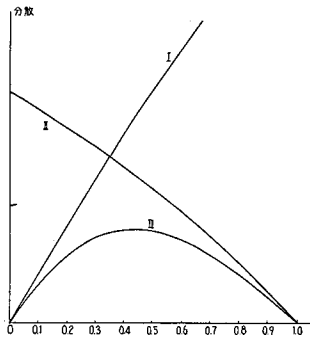
$$Y_2 = X_{22} + (X_{21} - X_{22})p_1$$

$$Y_3 = X_{32} + (X_{31} - X_{32})p_1$$

比較を具体的にするために、上述の数値例による計算結果を図に示すと、次のようになる。図中に、後に述べる最尤法による解の分散も入れてある。

まず、いわゆる分類の過誤をおかさない(1)の方法による推定量の分散は、分類の過誤をおかす(3)あるいは(4)の方法による分散よりも小さくはないことが明らかである。次に成分個体群 I の混合割合が約34%を境にして、(3)、(4)の方法による推定量の分散の大小関係が逆転する。分布の重複度という点からいえば(3)の場合が小さいにも拘らず、分散の大小関係が逆転することに注目していいであろう。

成分個体群 I の混合割合による分散の変化
 曲線 I : 場合(1)の p_1 及び場合(4)の p_1, p_2 の分散
 曲線 II : 場合(1)の p_1 及び場合(3)の p_1, p_2 の分散
 場合(2)の分散曲線は図の上方に出て終う。
 曲線 III : 最尤解の p_1, p_2 の分散



この例が具体的に示しているように、分類の過誤を避けることは、たとえそれが可能であっても、推定量の精度という点で余り意味はない。自然標識の場合、分布の重複していない縁辺の、従って比較的稀な事象を利用すれば、分類の過誤をおかさずに済すことは不可能ではないが、同様に余り意味のあることではない。

分類の過誤を生ぜしめる分布の重複度についていえば、それを α とすると、この例では、成分個体群頻度行列の行列式 Δ は、 $1 - \alpha$ となる。それは、分散に逆数として含まれており、分布の重複度 α の小さい標識を用いれば、それだけ Δ は大きく、推定量の分散も小さくなることが予想される。然しながら、分布の重複度だけでは、頻度行列が確定しないで、重複の仕方も関与している。更に、混合割合自身も関与している。この例は、ある頻度行列に対する評価が、混合割合如何で変ることを示している。

もし、製品検査の場合のように、ほぼ一定の混合割合が予想されるならば、標識の評価は、重複度と重複の仕方だけでできるであろう。このことは、混合割合の先験的確率分布として、ある値で1、その他で0という分布を想定して、分散の期待値(平均値)を求めていることに外ならない。従って、一般的に、混合割合の先験的確率分布が想定できるならば、平均としての評価ができよう。すなわち上例でいえば、先験的確率分布を $\phi(p)dp$ とすれば、

$$\bar{v} = \int_0^1 V(p) \phi(p) dp$$

を求めて、比較すればよい。

しかし、混合割合の先験的確率分布には、生態学的に、かつ数細に観察したときの魚群の混合の仕方と、調査の仕方、すなわち、標本抽出計画のあり方が関係していよう。その複雑さを考えれば、標識に固有の性質として、相互に評価することは非常に難しいのである。

労力、その他によって自ら限度はあるが、推定量の分散には、標本の大きさも関与している。従って、十分に大きい標本が、容易に抽出できるならば、標識に固有の性質として多少劣っていても、それを十分に補ない得ることがあるわけである。

III 標識の選択基準 (2) 統計的に確定した混合割合推定の可能性

前章の例で、成分個体群における観測標識の頻度分布が既知のときの特殊な解き方を述べている。それは観測標識の階級を、分類標識として排反する階級に再編し、その上で、得られた2元連立1次方程式を解いているわけである。すなわち、夫々、まず

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

という行列で座標変換し、更に、得られた2次の頻度行列の逆行列で、もう一度、座標変換している。それは、例えば、(3)の場合についていえば、

$$\frac{1}{\Delta_3} \begin{pmatrix} X_{32} & -X_{12}-X_{22} \\ -X_{31} & X_{11}+X_{21} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv \frac{1}{\Delta_3} \begin{pmatrix} X_{32} & X_{32} & -X_{12}-X_{22} \\ -X_{31} & -X_{31} & X_{11}+X_{21} \end{pmatrix} \equiv A$$

だから、Aという座標変換をすることに外ならない。それは、3行2列の頻度行列を変換して、2次の単位行列

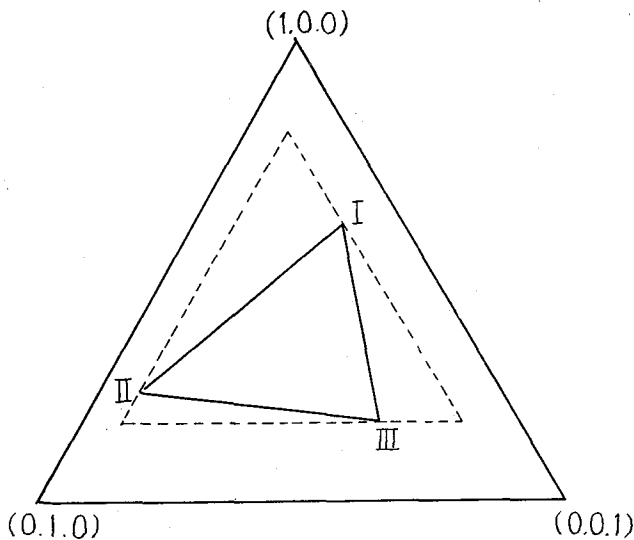
$$E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

になるような、2行3列の行列Aであれば、どんなAを用いても解き得るということである。そのようなAは、この場合、無数に存在するのである。例えば、前述の解を、その和が1になるように荷重して、加えると、それもまた、混合割合の不偏推定量になることは、明らかであろう。

この事情を次に考察するのであるが、最初に、現象的な点に触れておこう。一般的な頻度行列は、I行J列から成っているが、観測標識のある階級に関する各成分個体群の頻度 X_{ij} ($j=1, 2, \dots, J$) を相互に比較すると、最大値、最小値がある。それらを夫々、 $\max_j X_{ij}$; $\min_j X_{ij}$ とすると、それらの混合個体群の同じ階級の頻度 Y_i は

$$\max_j X_{ij} \geq Y_i \geq \min_j X_{ij} \quad (i=1, 2, \dots, I)$$

これは、標識階級が2のとき、ある個体群が、それらの混合個体群であるための必要かつ十分な条件であるが、階級数がそれ以上のときは、十分な条件ではない。各成分個体群の頻度分布は、 $\sum_i X_{ij} = 1$ の $(I-1)$



次元超平面上の点であり、ある条件を満足すれば、それらの点は、その上で、凸多角形を形成する。混合個体群の頻度分布を表わす点は、この多角形の内部になければならない。標識階級数が3のとき、図にかくと左図のようになる。すなわち、実線の $\triangle I II III$ が、成分個体群の形成する3角形であるが、点線の3角形は必要条件を満す点の集合である。 Y_i が必要かつ十分な条件を満すとき、例えば、算術平均値というような汎函数はどうなるかといえ、算術平均値 f は、標識階級に a_1, a_2, \dots という数値を与えて、 $f = \sum a_i Y_i$ という式で表わされ、この汎函数 f は、多角形の頂点で最

大値、最小値をとる。すなわち、混合個体群の数量化された観測標識の平均値は、成分個体群における平均値の最大値と最小値との間にある。例えば、成長に伴って平均脊椎骨数の変異の程度が次第に小さくなって来る傾向があるとき、抽出変動の問題もあるが、混合が次第に進行することを示していると考えられよう*。

さて、観測標識の階級数と成分個体群の数とが等しいとき、周知のように、解の存在する条件は、頻度行列の行列式が0でないということである。これは、各頻度分布が相互に独立なI次元ベクトルであるということであるが、このとき、成分個体群の頻度分布を表わす点は、(I-1)次元の超平面 $\sum X_i = 1$ の上で、J (I=J)ヶの頂点をもつ凸多角形を形成している。行列式が0になることは、Jヶの頂点をもつ凸多角形を形成しない。例えば、標識階級数が3のとき、3点が一直線上にのるということである。極端な場合、2点、あるいは3点が一致することもある。

このことは、一般に、頂点を形成する個体群を成分として、少くとも一つの成分個体群が、それらの混合個体群として表わされるということである。このように、与えられたI (=J)次の頻度行列の行列式が0であって、その行列の中に、0でない γ (<I)次の小行列式が見出され、より高次の小行列式がすべて0ならば(このことを頻度行列の階級数が γ であるという)、適当な変換によって、頻度行列は、 γ 行J列($\gamma < J$)の頻度行列に変換され、観測標識の階級数が成分個体群の数より少ない場合と同等になる。

$\gamma = 1$ のとき、不能であるが、実際問題として、頻度分布が同じであるような標識で問題をとり上げることはあり得ないから、 $\gamma \geq 2$ である。そして $\gamma < J$ ならば解は不定、即ち、ある混合個体群の混合割合が一意的に確定しない。

この場合、即ち、観測標識の階級数Iが成分個体群の数Jより小さく、頻度行列の階級数がIのとき、次のようにして一般解が求められる。行列式が0でない適当なI次の正方行列(X'_{ij})に注目して、頻度行列(X_{ij})を(X'_{ij} , X_{ij-1})に分け($|X'_{ij}| \neq 0$), $X'_{ij} p'_i = Y_i$ ($j=1, 2, \dots, I$)を解いて、その解を特殊解 p'_i のI個の元とし、(特殊解は、このI個以外の元は0とする)次に、

$$\begin{pmatrix} X'^{-1}_{ij} & X_{i,j-1} \\ -E_{j-1} \end{pmatrix} \quad E_{j-1} \text{は } (J-1) \text{ 次の単位行列}$$

を求めると、一般解は次のようになる**。

$$\begin{pmatrix} p_1 \\ \vdots \\ p_i \\ p_{i+1} \\ \vdots \\ p_j \end{pmatrix} = \begin{pmatrix} X'^{-1}_{ij} Y_i \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{j-1 > 0} \lambda_{j-1} \begin{pmatrix} X'^{-1}_{ij} & X_{i,j-1} \\ \vdots & \vdots \\ -E_{j-1} \end{pmatrix}$$

このように、一般解には(J-1)個のパラメーターが荷重として含まれている。しかし、解が混合割合であるためには、 $1 \geq p \geq 0$ でなければならないから、パラメーター λ_{j-1} の動く範囲は定まっている。ここで、抽出変動を伴うのは、成分個体群の母頻度分布は既知として、特殊解の部分だけである。そして、それは、変量の一次関数だから、普通の公式に従って、この不確定関係に対する分散を求めることができよう。

残る場合は、観測標識の階級数が成分個体群の数より多い場合である。この場合にも、I行J列(I>J)の与えられた頻度行列の階級数が、Jより小さければ、前の場合に帰着する。I行J列の行列では、その階級数は $\min(I, J)$ より大きくなり得ないから、残るのは、少なくとも一つのJ次の小行列式が0でない場合である。このとき、そのJ次の小行列の部分を用いて、解は確定する。0でないJ次の小行列式が幾つかある場合には、前章の例のように解き方は幾通りもあり、見掛上相異しても、確定した混合割合の不偏推定量を与えるのである。

以上から明らかのように、混合割合の推定が一意的に可能であるためには、少なくとも、観測標識の階級数が、成分個体群の数より少なくはないことが必要である。これが精度とは別の標識選択の基準であることはいうまでもない。頻度分布の重複に由来する精度の問題は、標本の大きさを増すことによって改善するこ

* Widrig and Taft (2b) は、この現象を利用してゐる。

** この解の数値例はVI論議に詳しい。

とができる。しかし、標識階級数の不足に由来する構造的な不確定さは、それによって解決しようのないものである。この不確定さ如何によっては、推定量の精度を良くする努力は、無意味になることがあるであろう。

このことは、後に具体的に論議するように、そのような標識による分析を、全く無意味だとすることは勿論ない。それは混合割合の推定値に、可能な最大値、最小値を与えるであろう。これは、信頼区間などの区間推定とは違って、その間のどの値がいかにして何の確率的評価を与えるものではないが、生物学的考察に際して有用な知識を与えるであろう。

この不確定さを根本的に解決するには、例えば、脊椎骨数などのように、階級数が固定されているならば新たな標識を組合せて、複合した標識による外はない、標本の大きさによって自ら限度はあるが、連続的変量と見做される標識は、階級の分け方に条件がないから、それだけ効用が大きいといえよう。

IV 最尤法による確定的な混合割合の推定

前述のように、観測標識の階級数 I が成分個体群の数 J より大きいとき、(勿論頻度行列の階級数は J として) 例に示したように、解き方は一通りではなく、それぞれに精度も異なっているから、当然精度の良い解き方が問題になる。例に示した解き方を省みると、ある標識階級を無視したり、あるいは、標識階級を組合せることによって、階級間の差を埋めて終っており、直観的にいって、標本の提供する情報を部分的にしか利用していないことになる。

Fisher, R. A. は、推定すべき母数 a の尤度を $L(x, a)$ とするとき、 $\frac{d}{da} \log L$ の分散を、大きさ n の標本 x によって提供される a に関する情報と定義し、 a の函数 $T(a)$ のどんな不偏推定量でも、その分散は $T'(a)^2/H$ (H : Information) より小さくならないことが示されている。ここでは、 $T(a) = a$ だから、この分散の下限は $1/H$ である。前述の例では、この下限は、

$$H(p_1) = n \sum_{i=1}^3 \frac{(X_{i1} - X_{i2})^2}{Y_i}$$

の逆数であり、これが混合割合 p_1 によってどのような値をとるかは、 $nv = n/H(p_1)$ として図に示されている。第II章で例示した解き方が、標本の与える情報を、どのように利用していないか明らかであろう。

最尤法による解は、 n が十分に大きいとき、分散として $1/H(p)$ をもつような解、即ち混合割合の推定量を与えるのである。この場合、尤度は、標本度数を n_i ($\sum n_i = N$) として、

$$L = \prod_{i=1}^I Y_i^{n_i} \quad Y_i = \sum X_{ij} p_j \quad (\sum X_{ii} = \sum Y_i = \sum p_i = 1)$$

であるが、これを最大にするように p_j を求めることは、 J が少し大きくても、また、 $J = 2 \sim 3$ としても、混合割合の異なる個体群、従って標本の数が多くなれば面倒である。そこで、Method of Scoring と呼ばれる逐次近似法によるのが、機械化されて实际的であろう。原理は成書(例えば、Rao, C. R. Advanced Statistical Methods in Biometric Research; p. 165)にあるので、ここでは、計算に必要な式を示すに止める。

近似値として、 $0 \leq p_i$ が与えられたときその値における、 p_j の Efficient score を S_j 、また情報を $H_{jj'}$ とすると

$$S_j = \sum \frac{n_i}{Y_i} (X_{ij} - X_{ij}) ; H_{jj'} = \sum \frac{N}{Y_i} (X_{ij} - X_{ij})(X_{ij'} - X_{ij'}) \quad (j, j' = 1, 2, \dots, J-1)$$

で計算される。 $0 \leq p_j$ に対する修正値 dp_j とすると、 dp_j は次の連立一次方程式を解いて得られる。

$$(I) \begin{cases} H_{11} dp_1 + H_{12} dp_2 + \dots + H_{1, J-1} dp_{J-1} = S_1 \\ H_{21} dp_1 + H_{22} dp_2 + \dots + H_{2, J-1} dp_{J-1} = S_2 \\ \dots \\ \dots \\ H_{J-1, 1} dp_1 + H_{J-1, 2} dp_2 + \dots + H_{J-1, J-1} dp_{J-1} = S_{J-1} \end{cases}$$

この連立方程式を解くことは、 $S_j, H_{jj'}$ を分解して、

$$S_j = S'_j - S'_j; H_{jj'} = H'_{jj'} - H'_{j'j} - H'_{jj} + H_{jj}$$

とすると、次の連立方程式を解くことと同じである。

$$(2) \begin{cases} H'_{11}dp_1 + H'_{12}dp_2 + \dots + H'_{1J}dp_J + M = S'_1 \\ H'_{21}dp_1 + H'_{22}dp_2 + \dots + H'_{2J}dp_J + M = S'_2 \\ \dots \\ H'_{J1}dp_1 + H'_{J2}dp_2 + \dots + H'_{JJ}dp_J + M = S'_J \\ dp_1 + dp_2 + \dots + dp_J = 0 \end{cases}$$

dp_j が得られれば、 $dp_j = {}_0p_j + dp_j$ を第1近似として、同様の計算をする。十分な近似が得られるまで繰返し最終値が得られたとき、(1)の場合は、その値における行列式 $|H_{jj'}|$ の H'_{jj} の余因子が、 \hat{p}_j の分散を与え、それから \hat{p}_j の分散も計算できる。(2)の場合は、そのときの行列

$$\begin{pmatrix} H'_{jj} & 1 \\ 1 & 0 \end{pmatrix}$$

の逆行列で、 (H'_{jj}) に対応する部分が、 \hat{p}_j ($j = 1, 2, \dots, J$) の分散行列になっている。

計算を始める最初の値は、頻度行列で、適当な0でないJ次の小行列式の部分を使って求められよう。便宜のためにこの計算法の数値列を次に示しておこう。

最尤解計算法の数値列

第II章の例のような頻度行列が与えられており、大きさ100の標本で、標本度数が $n_1=76$, $r_2=5$, $n_3=19$ だったとしよう。先づ ${}_0p_j$ を、階級2, 3を合せて求めよう。既述のように方法は幾通りもあるが、ここでは、例の4通りの中でも最も精度の悪い組合せを、説明のためにとる。

$$\begin{pmatrix} 0.85 & 0.75 \\ 0.15 & 0.25 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0.76 \\ 0.24 \end{pmatrix}$$

を解いて ${}_0p_1=0.10$, ${}_0p_2=0.90$ 。これを使って、階級頻度 Y_i 、分解した score information を求める。

	X_{i1}	X_{i2}	n_i	Y_i	$\frac{1}{Y_i}X_{i1}$	$\frac{1}{Y_i}X_{i2}$
階級1	0.85	0.75	76	0.760	1.118	0.987
階級2	0.15	0	5	0.015	10.000	0
階級3	0	0.25	19	0.225	0	1.111
			100	1.000		

	$S'_1 = \frac{n_i}{Y_i}X_{i1}$	$S'_2 = \frac{n_i}{Y_i}X_{i2}$	$H'_{11} = \frac{N}{Y_j}X_{i1}^2$	$H'_{12} = \frac{N}{Y_i}X_{i1}X_{i2}$	$H'_{22} = \frac{N}{Y_i}X_{i2}^2$
階級1	84.968	75.012	$N \times 0.950$	$N \times 0.839$	$N \times 0.740$
階級2	50.000	0	$N \times 1.500$	$N \times 0$	$N \times 0$
階級3	0	21.109	$N \times 0$	$N \times 0$	$N \times 0.278$
	134.968	96.121	$N \times 2.450$	$N \times 0.839$	$N \times 1.018$

この場合、(1)で解けばすぐに出るが、(2)の方法ですと、次の行列の逆行列を求めることになる。

$$A = \begin{pmatrix} 245.0 & 83.9 & 1 \\ 83.9 & 101.8 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

逆行列 A^{-1} を求めると

$$A^{-1} = \begin{pmatrix} 0.558 \times 10^{-2} & -0.558 \times 10^{-2} & 0.100 \\ -0.558 \times 10^{-2} & 0.558 \times 10^{-2} & 0.900 \\ 0.100 & 0.900 & -100.000 \end{pmatrix}$$

従って dp_j は

$$dp_1 = 0.558 \times 10^{-2} \times 134.968 - 0.558 \times 10^{-2} \times 96.121 = 0.218$$

$$dp_2 = -0.558 \times 10^{-2} \times 134.968 + 0.558 \times 10^{-2} \times 96.121 = -0.218$$

第1近似として

$${}_1p_1 = 0.10 + 0.218 = 0.318$$

$${}_1p_2 = 0.90 - 0.218 = 0.682$$

同様の計算を繰返して

$${}_2p_1 = 0.318 - 0.032 = 0.286$$

$${}_2p_2 = 0.682 + 0.032 = 0.714$$

更に、

$${}_3p_1 = 0.286 + 0.005 = 0.291$$

$${}_3p_2 = 0.714 - 0.005 = 0.709$$

ここで、近似を止めて、この値における A^{-1} を求めると、

$$A^{-1} = \begin{pmatrix} 1.142 \times 10^{-2} & -1.142 \times 10^{-2} & 0.291 \\ -1.142 \times 10^{-2} & 1.142 \times 10^{-2} & 0.709 \\ 0.291 & 0.709 & 100.000 \end{pmatrix}$$

$$\text{即ち、} \hat{p}_1 = 0.291 \quad \hat{p}_2 = 0.709$$

$$v(\hat{p}_1) = v(\hat{p}_2) = 0.01142$$

ということになる。因みに、 ${}_0\hat{p}_1 = 0.10$, ${}_0\hat{p}_2 = 0.90$ の分散は 0.1824 である。

V 適合度その他の仮説検定と成分個体群に関する不確かさに由来する推定混合割合の不確か性

成分個体群に関して、恰も十分に既知であるかのように考察して来たが、現実にはほとんど常にそうでないといっているであろう。

(1) 成分個体群の標識頻度分布が、母分布でなく標本頻度分布であるという不確かさが、ほとんど常に存在しよう。しかし、この場合、成分個体群からの J 個の標本があるならば、それらに関しては、混合割合 p_j のどれかが 1 であり、その他が 0 であることは確かである。そして、頻度分布 X_{ij} をも推定すべきパラメターとして、最尤解を求めることができる。

この場合、 X_{ij} に関する情報は、成分個体群からの標本ばかりでなく、混合個体群の標本からも来るわけである。そこでもし、混合個体群からの標本が非常に大きい場合には、それから得られる情報の方が大きいことがあり得よう。

この事情を考慮すると、ここで重要なのは、頻度分布が母分布でないというような不確かさよりも、それが成分個体群からの標本であるという確かさである。この確かさ、即ち、成分個体群が実践的に定義されているか否かという確かさは、非常に難しい問題である。この確かさには、既往の生態学的知識が大きく関与していることは明らかであろう。サケ・マスの場合、後に詳しく論ずるが、見易いように思われるのは、その母川回帰の性質に基づいて、代表的標本抽出の計画が、かなり明確に立案し、実行され得るという可能性の故である。

(2) 実践的にも保証され、かつ、すでに考慮している J 個の成分個体群で十分であるかどうか、他に有意に関与している成分個体群が存在しないかどうかなどの不確かさがある。この問題は、その背景によってさまざまであるが、ここでは先ずある分析結果の考察に関連した問題を考える。

最尤法による解は、確かに、 $\sum p_j = 1$ という条件を満たしている。しかし、 $p_j \geq 0$ という条件を満たしているとは限らない^{*}。それは、その標本点が、頻度行列で定まる $(I-1)$ 次元の超平面上の J 多角形の外にあるときである。便宜のために、 $I = 4$, $J = 3$ の場合を考え、例えば $p_j > 0$, $\sum p_j = 1$ で次のような解があったとしよう。

$$Y = p_1 X_1 - p_2 X_2 + p_3 X_3$$

移項して、 $p_1 + p_3 = 1 + p_2$ であることを考慮すれば、新たに、 $a_0 + a_2 = a_1 + a_3 = 1$, $a_j > 0$ ($j = 0, 1, 2, 3$) として

³⁾
* $p_j \geq 0$ という条件を入れて解くこともできる。しかし、それは模型の妥当性について確証がなくてのことであろう。

$$a_0 Y + a_2 X_2 = a_1 X_1 + a_3 X_3 \quad (\equiv Z)$$

即ち、Zは、YとX₂、X₁とX₃を結ぶ線分の交点を示している。このようにして、YとX₁、X₂、X₃との相互位置は明らかにされるが、その上で、適当な仮説検定、この例では p₂ = 0 というような仮説の検定を尤度比検定法を利用してできよう。それは、m個のパラメーターを仮説としてそれぞれある特定の値に固定したときの最大尤度を L₁、そうしないときの最大尤度を L₀ とすると、 $\lambda = L_1/L_0$ として $-2 \log \lambda$ が自由度 m の X² 分布をすることを利用するのである。

次に、このような仮説が否定される標本は除いて、例えば、次のような X² 検定で適合度を検定できよう。即ち、情報行列式 |H_{ij}| の H_{ij} の余因子を H^{ij}、Efficient score を S_i とするとき、

$$X^2 = \sum \sum H^{ij} S_i S_j$$

は (i) 成分個体群の頻度分布が既知ならば、K(I - J)

(ii) 成分個体群からの標本抽出が確かならば、(K - J)(I - J)

という自由度の X² 分布をすることを利用するのである。

ここで重要なことは、標識の階級数と成分個体群の数とが等しいときは、統計的に模型の妥当性を検討する道がないこと、言い換えれば、生物学的に模型の妥当性が保証されていないということである。製品の規格検査の場合、検査のプロセスが模型の妥当性を保証していることは明らかであろう。

(3) このようにして、新たに成分個体群を付け加える必要があるとき、それを推定する問題を考えよう。この場合、J個の成分個体群に関しては、実践的な保証があるわけであるが、付加される個体群については、それが無いことになる。そして、極端な場合として、成分個体群に関して全くにも実践的保証がない。即ち、頻度分布も、成分個体群の数も分らない場合があり得る。心理的にこれに近い場合は確かにあるが、もし字義通りだとすれば、そのような事情の下で混合問題を想定すること自身、余りに観念的だと言わざるを得ない。結論から言えば、この場合、解は不確定である。次にその事情を考察しよう。

この場合、各標本の大きさが十分に大きいとき、即ち、模型として解かなければならない方程式群は次のようになる。

$$(p_{kj})(X_{ji}) = (Y_{ki})$$

ここに、(p_{kj})は、混合割合を行ベクトルとする K 行 J 列の行列、(X_{ji})は頻度行列を転置した J 行 I 列の行列、(Y_{ki})は、混合個体群の頻度分布を行ベクトルとする K 行 I 列の行列である。

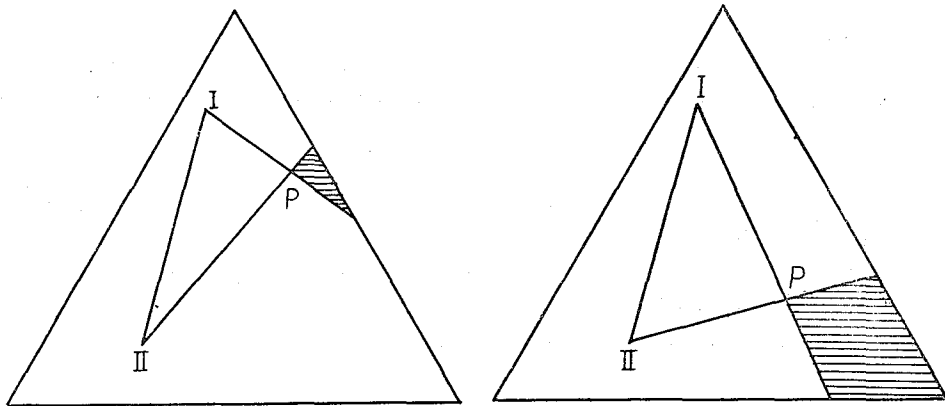
さて、K > I として、行列 (Y_{ki}) の階数は高々 I である。従って K 組のうち、適当に I 組を選べば、それ以外の組の方程式群は、この I 組の方程式群から、その一次結合として導かれる。即ち、この方程式群は、解く上で、この I 組の方程式群と同等である。次に、選ばれた I 組の方程式群について、混合割合行列は、I 行 J 列であるが、I > J として、その階数は高々 J である。J 組の方程式群を適当に選ぶと、その混合割合行列の逆行列は存在するから、この J 組の方程式群を頻度行列に関して解くことができ、それを残る (I - J) 組方程式群に代入して頻度行列を消去できる。ところが、この (I - J) 組の方程式群における混合割合ベクトルは、J 個の混合割合ベクトルの一次結合で表わされる。従って結局、この (I - J) 組の方程式群の左辺は、選んだ J 個の混合個体群頻度分布ベクトルの一次結合になり、最初選んだ K 行 I 列の行列 (Y_{ki}) の階級が I であるという仮定と矛盾する。

I ≤ J のときも同様に不確定であるが、I > J の場合と不確定さの事情が若干相異なるのである。I ≥ J、K > J ならば、少なくとも、頻度行列を消去する可能性、即ち (K - J) 個の混合割合を J 個の混合割合だけで表現する可能性がある。しかし、I < J ならば、その可能性がないわけである。I ≥ J のとき、J 個の混合割合の一次結合で表現されるその結合の係数は、上述からも明らかなように、(K - J) 個の混合個体群の頻度分布が、J 個の混合個体群の頻度分布の一次結合で表わされるときその係数に外ならない。I = J ということは、この係数が任意に定められるということであり、しかも、どの場合も尤度は同じである。I > J という条件は、尤度を最大にするような係数が定められるという条件である。このように、係数が一意的に定まり、混合割合が頻度行列に係りなく、相互にある関係にあるということは、生物

学的考察を進める上で如何に有用かは論ずるまでもないことであろう。

この事情は、J個の個体群に関して、その混合割合が既知であることの重要性を示している。実際的には、このJ個については、それらが成分個体群であることを確認するのが実践的に最も容易であるわけである。いうまでもなく、J個の成分個体群のうち、若干のものについて実践的に確かならば、それだけ、不確定さは減少するであろう。もし、実践的に不確かだとしても、実質的にそれに近い標本が含まれているならば、矢張り、それだけ、不確定さは減少するであろう。不確定さは何れにしても残るけれども、調査計画如何によって、それを小さくし得るわけである。

この事情を図で説明すると、例えば、標識階級数が3で、成分個体群が2つ既知としよう。適合しない十分大きい標本をPとすると、その位置によって不確定さの程度がどのように相異なるかは、明らかであろう。



即ち、斜線の部分が、3番目の成分個体群の存在する可能性のある範囲である。

IV 論 議 北洋におけるアジア系、北米系ベニザケの混合に関する諸研究の方法論的論議と推定混合割合の不確定性の分析

北洋のサケ、マス混合に関する調査研究は、混合の分析を課題とする調査研究としては、恐らく最初の大規模なものと考えられるが、この問題に関連して、混合個体群の分析に関する総括的な論議を試みよう。

共同調査の初期には、当然のことながら、北米系、アジア系サケ、マスが、どのような標識によって識別され得るか、或いは、より良く識別出来るか、即ち、標識の探究に、努力が向けられていたと言えるだろう。その結果として用いられている標識は、その主要なものを挙げると、次のようである。

(1) 体節の形質 ベニザケについて、側線上の鱗数その他7形質に基づく一次判別函数(Fukuhara⁴⁾⁵⁾。

(2) 鱗の成長特性 若干の相違はあるが、淡水生活期を示す年輪半径 circuli 数など (Shepard, Kobayashi¹⁾²⁾³⁾⁸⁾。

(3) 寄生虫学的標識 条虫 *Triacnophorus crassus* の幼生の寄生 (Margolis⁹⁾¹⁰⁾¹¹⁾。重点はベニザケにおかれている。

(4) 血清学的標識 或る種の抗原の有無 (Ridgway et al.¹³⁾¹⁴⁾。これもベニザケについてである。

この外に、年々、公海において、相当の規模の標識放流が行なわれ、概括的な考察には年齢組成なども使われている。

この段階で採られている方法は、可能な成分個体群として、可成り明確に実践的に定義し得る個体群について、即ち、河川標本について、標識頻度分布を比較することである。その場合の基準は、分布の重複度であり、それと同様に、距離、或いは、単純な場合には標識平均値の差である。その代表的な例は、体節の形質による研究に見出される。同等でない標識を付加することによって、一般化された距離は増加する。そし

て次の段階として、便宜のために、距離を最大にするような一次判別函数を標識として採用しているのである。この基準が直観的なものであり、間違っているわけではないが、それだけで十分でないことは、既に論じた通りである。このように、観測標識として一次判別函数を採用した後に、分類標識としては、2つの成分個体群平均値の中央値によって、標本空間を2分している。従って、この2分された観測標識階級の頻度は、混合割合としては偏りを避け得ないから、この論文の模型に従って、改めて不偏推定量を求めるわけである。

この段階の研究で、分散分析、有意性の検定などが行なわれているが、その視点は、所謂種族分析に適用されている場合と異っている。所謂種族分析では、対象個体群の異質性から一步進めてその生物学的内容を推論することを目標としている。異質性を言うためには、標本の大きさを十分大きくすれば可能であるが、推論を進めるために、更に種族の形質とか、形質の遺伝性とかが論議されることが往々ある。しかし、ここでは、異質性は明らかであり、また混合割合の推定には、標識としての形質が、このような性格を持つことは、原理的に少しも必要ではない。或る個体群がそれによって成分個体群として定義されればよいのである。有意性の検定は、適当な大きさの標本で、有意になることを基準にして、距離の大きい形質を選ぶために行なわれている。

条約は、確かに、北米系サケ・マスとアジア系サケ・マスの混合を問題にしている。また、確かに、北米の河川で発生したサケ・マス或いは、北米の河川に産卵回帰するサケ・マスという定義も無意味ではない。しかし、問題は、成分個体群として、そのような個体群を定義し得るかということである。寄生虫学的標識および血清学的標識は、観測標識として階級数は2つしかない。また、体節的形質による一次判別函数は、或る定義された北米系個体群とアジア系個体群の距離を最大にするように定められている。従って、成分個体群としてそのような定義が出来ないとすれば、既述のように、これらの標識は、適格性を原理的には欠くわけである。

現在、北太平洋には、ベニザケの大きいストックは、アジア側ではカムチャッカ半島の南端、北米側では Bristol 湾、フレイザー水系の3地域にあって、問題の水域で混合しているのは、前2者だけと考えられている。そして、アジア側では、2河川、Bristol 湾では5河川あって、それぞれの回帰群で、標識頻度分布は相違している。

そこで、それらの諸研究では、例えば、北米系成分個体群の標識頻度分布は、各河川群の頻度分布を、それぞれの回帰量を荷重にして荷重平均したもので定義されている。それは、北米系ベニザケの標識頻度分布としては、確かに合理的である。しかし、成分個体群の定義としては、合理性を欠くのである。即ち、この定義は、各河川回帰群が、同じ割合で混合しながら、考える回遊期間中行動していることを前提としているからである。この前提が如何に偏ったものであるかは論ずるまでもないであろう。事実、標識放流再捕の結果は、この前提の成立しないことを示している。従って、現在使われている標識およびその方法は、既述のように、構造的な欠陥を原理的にもっていると言わなくてはならない。それは、勿論、無意味だということではない。

寄生虫学的標識は、分布範囲などの推論を考えれば、確かに有力な標識には違いないが、或る一種の寄生虫の有無を標識とする限り、その階級数は2である。血清学的標識は、原理的に可能であるかも知れないが、現在使われているように、階級数が2に、技術的に固定されているならば、前者と同様であろう。これらの標識は、同種、或いは異種の標識を組合せることによってしか、階級数を増し得ない。これに反して、体節的形質の一次判別函数は、連続的変量として前提すれば、或る程度の任意性を持っている。定義されたような2個の成分個体群の距離を最大にするように定められた一次判別函数よりも、幾つかの成分個体群間の平均距離を最大にするように定められた一次判別函数の方が精度が良いことが予想されるが、精度の問題には距離だけが関与しているのではないから、そのまま使うことも不合理ではないであろう。

ここで、これらの標識および方法による構造的な不確定さを、少し詳しく考察しておこう。1959年の Fukuhara et al の報告第16表に、判別函数値の空間を分類標識として2分したときの階級頻度が示されてい

る。北米系として、論議を簡明にするために、ブリストル湾の5河川回帰群を採り、(コディアク、チグニクを除くわけである)、またアジア系としては、Fukuhara et al. も採用しているものをを用いよう。アジア系成分個体群には種々問題はあるが、河川における標本のない現在、止むを得ないものである。

それから、ブリストル湾の各回帰群、アジア系成分個体群の標識頻度分布は、次のように得られる。

成分個体群番号	アジア系		ブリストル湾系			
	(1)	(3)	ナクネク河 (4)	エゲシク河 (5)	ウガシク河 (2)	クヴィチャク河 ウッドリバー (6)
階級I	0.763	0.206	0.189	0.223	0.152	0.408
階級II	0.237	0.794	0.811	0.777	0.848	0.592
標本の大きさ	581	508	180	184	184	164

用いられた標本の大きさも考えて、一応、これらの頻度分布を母分布と見做すことにする。既述の方法に従って特殊解を求めるのに、どの2つの成分個体群を採ってもいいが、重複度の少ないアジア系群とクヴィチャク河群を採ろう。

$$\begin{pmatrix} 0.763 & 0.152 \\ 0.237 & 0.848 \end{pmatrix} - 1 \equiv \begin{pmatrix} 1.389 & -0.249 \\ -0.389 & 1.249 \end{pmatrix}$$

だから、或る混合個体群の観測頻度分布を Y_1, Y_2 として、特殊解は、

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = \begin{pmatrix} 1.389Y_1 - 0.249Y_2 \\ -0.389Y_1 + 1.249Y_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

一般解は、

$$\begin{pmatrix} 1.389 & -0.249 \\ -0.389 & 1.249 \end{pmatrix} \begin{pmatrix} 0.206 & 0.189 & 0.223 & 0.408 \\ 0.794 & 0.811 & 0.777 & 0.592 \end{pmatrix} = \begin{pmatrix} 0.088 & 0.030 & 0.116 & 0.420 \\ 0.912 & 0.970 & 0.884 & 0.580 \end{pmatrix}$$

だから、流動パラメーターを正になるように定めて、

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = \begin{pmatrix} 1.389Y_1 - 0.249Y_2 \\ -0.389Y_1 + 1.249Y_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -0.088 \\ -0.912 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -0.030 \\ -0.970 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -0.116 \\ -0.884 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} -0.420 \\ -0.580 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

になる。北米系の混合割合というのは、言うまでもなく、 $p_2 + p_3 + p_4 + p_5 + p_6$ であり、この場合 $1 - p_1$ に等しい。特殊解の部分は、標本によって定まって終うから、例えば、北米系の混合割合に関する構造的な不確定さの程度は $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ が条件

$$1 \geq 1.389Y_1 - 0.249Y_2 - (0.088\lambda_1 + 0.030\lambda_2 + 0.116\lambda_3 + 0.420\lambda_4) \geq 0$$

$$1 \geq -0.389Y_1 + 1.249Y_2 - (0.912\lambda_1 + 0.970\lambda_2 + 0.884\lambda_3 + 0.580\lambda_4) \geq 0$$

$$1 \geq \lambda_1 \geq 0; 1 \geq \lambda_2 \geq 0; 1 \geq \lambda_3 \geq 0; 1 \geq \lambda_4 \geq 0$$

を満足するときの $f = 0.088\lambda_1 + 0.030\lambda_2 + 0.116\lambda_3 + 0.420\lambda_4$ の値によって定まる。^{*}

この場合、先ず明らかなのは、 $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ で f は最小値0になる。これは、距離の最も遠い成分個体群を選んでいるからであるが、このことから、北米系の混合割合は少なく見積っても $-0.389Y_1 +$

* 多数の標本について計算を行なうには、一般解を求めたのち、Simplex 法によって、高速計算機で解き得よう。

1. $249Y_2$ より小さくはない、或いは、アジア系の混合割合は、高々 $1.389Y_1 - 0.249Y_2$ であるという推論が可能である。

次に最大値であるが、

$$-0.389Y_1 + 1.249Y_2 - (0.912\lambda_1 + 0.970\lambda_2 + 0.884\lambda_3 + 0.580\lambda_4) \geq 0$$

$$1 \geq \lambda_1 \geq 0; 1 \geq \lambda_2 \geq 0; 1 \geq \lambda_3 \geq 0; 1 \geq \lambda_4 \geq 0$$

は、4次元の5面体であるが、(i)これを、4次元の超平面 $1.389Y_1 - 0.249Y_2 - (0.088\lambda_1 + 0.030\lambda_2 + 0.116\lambda_3 + 0.420\lambda_4) = 0$ が切断するならばその切断面上の $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ の組に対して、 f は最大値 $1.389Y_1 - 0.249Y_2$ を採る。(ii)若し、切断しないならば、原点を除く多面体の頂点のどれかで最大値を採り、そのときの f は、 $1.389Y_1 - 0.249Y_2$ より小さい。

数値例を示す。1960年の Fukuara の報告第10表から、次の4例を選んだ。

	抽出位置	年月日	標本の大きさ	観測頻度	
				階級 I	階級 II
例 1	168°15' E; 50°10' N	16/V/'58	117	0.556	0.444
例 2	175°00' W; 51°30' N	8/V/'58	70	0.471	0.529
例 3	170°02' W; 53°30' N	17/V/'58	50	0.320	0.680
例 4	155°00' W; 55°00' N	15/V/'58	88	0.386	0.614

特殊解は、0の元を除いて列ベクトルで表わすと、それぞれ

例 1	例 2	例 3	例 4
0.662	0.522	0.275	0.383
0.338	0.478	0.725	0.617

であり、例1、例2では切断面はなく、例3、例4で切断面がある。5面体の原点を除く頂点は、列ベクトルで表わして、例1の場合

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \begin{pmatrix} 0.371 & 0 & 0 & 0 \\ 0 & 0.348 & 0 & 0 \\ 0 & 0 & 0.382 & 0 \\ 0 & 0 & 0 & 0.583 \end{pmatrix}$$

従って、 f の最大値は、 $\lambda_1 = \lambda_2 = \lambda_3 = 0; \lambda_4 = 0.583$ で0.240となる。即ち、例1では、北米系の混合割合は、大きく見積って0.578、小さく見積って0.333である。例2で f の最大値は、同じ頂点で0.347になる。

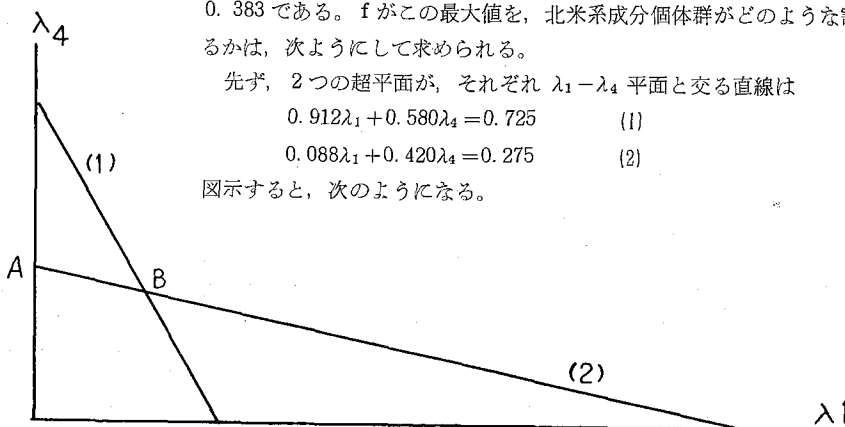
切断面がある例3、例4の場合、 f の最大値は、既に、特殊解で求められている。例3で0.275、例4で0.383である。 f がこの最大値を、北米系成分個体群がどのような割合のときにとるかは、次ようにして求められる。

先ず、2つの超平面が、それぞれ $\lambda_1 - \lambda_4$ 平面と交る直線は

$$0.912\lambda_1 + 0.580\lambda_4 = 0.725 \quad (1)$$

$$0.088\lambda_1 + 0.420\lambda_4 = 0.275 \quad (2)$$

図示すると、次のようになる。



$f < 0.275$ の直線は、(2)より下方にあって、(2)に平行である。(2)と λ_4 軸の交点、(1)と(2)の交点をそれぞれ A, B とすると、

$$A = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.655 \end{pmatrix}, \quad B = \begin{pmatrix} 0.437 \\ 0 \\ 0 \\ 0.563 \end{pmatrix}$$

同様に、2つの超平面が、 $\lambda_2 - \lambda_4$ 平面と交わる2直線の交点を C, D とすると、

$$C = \begin{pmatrix} 0 \\ 0.372 \\ 0 \\ 0.628 \end{pmatrix}, \quad D = \begin{pmatrix} 0 \\ 0 \\ 0.477 \\ 0.523 \end{pmatrix}$$

になる。前図からも明らかのように、線分 AB 上の点は、 f を最大値 0.275 にする。従って、 $1 \geq a_i \geq 0$, $\sum a_i = 1$ ($i = 1, 2, 3, 4$) とし、

$$a_1 A + a_2 B + a_3 C + a_4 D$$

で表わされる点は、 f を最大にする。各成分個体群としてとり得る最大値は、ナクネク群で 0.795, エゲジク群で 0.747, ウガシク群で 0.820, クヴィチャク群で 0.725, ウッド・リバー群で 0.655 であるが、そのとき北米系として最大の混合割合をとるわけではない。この条件を入れると、ナクネク群で 0.437, エゲジク群で 0.372, ウガシク群で 0.477, クヴィチャク群で 0.345, ウッドリバー群で 0.655 ということになる。勿論、最小値は 0 である。

同様に、例 4 では、A, B, G, D は次のような点になる。

$$A = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.912 \end{pmatrix}, \quad B = \begin{pmatrix} 0.111 \\ 0 \\ 0 \\ 0.839 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ 0.095 \\ 0 \\ 0.905 \end{pmatrix}, \quad D = \begin{pmatrix} 0 \\ 0 \\ 0.122 \\ 0.878 \end{pmatrix}$$

以上述べたような解釈の多様性が、この例では標本空間を 2 分して推定する方法の構造的な不確かさである。Fukuhara et al. は、この多様な解釈の中の一つを採っているだけである。そのような根拠は、分析自身にはない。ここで最大値、最小値に言及しているが、この意味でそれは信頼限界などの統計的区間推定とは、本質的に相違するのである。従って、Fukuhara et al. は Worlund¹⁶⁾ の方法で信頼限界を設定しているが、それでこの不確かさを補えるわけではない。標本の大きさを如何に大きくしても残るのである。

1959年の Margolis の報告によれば、1958年のベニザケ成魚に見られる *Triaenophorus* の寄生率は、アジア系とブリストル湾の 5 河川群とは、次のように推定されている。

	アジア系	ナクネク	エゲジク	ウガシク	クヴィチャク	ウッドリバー
非寄生	1	0.981	1	1	1	0.440
寄生	0	0.019	0	0	0	0.560

ここ数年の調査研究によって、ウッドリバー群およびナクネク群以外に *Triaenophorus* は寄生していないのが、ほぼ確かであり、従って、この表からただちに明らかなることは、アジア系、エゲジク、ウガシク、クヴィチャク群は、この標識では区別出来ないことである。このことは、どんな混合個体群も、それがブリストル湾系 5 群の混合個体群と解釈し得るということである。これは自明のことであるが、そこで、この 4 群の任意の混合個体群を、成分個体群 A としておこう。この成分個体群とウッドリバー群から、特殊解を求め、 $1 \geq \lambda_1 \geq 0$ とし一般解は次のようになる。 P_A は混合個体群であるが、いま成分個体群 A と定義したものの混合割合、 P_2, P_6 は前例と同じである。

$$\begin{pmatrix} P_A \\ P_6 \\ P_2 \end{pmatrix} = \begin{pmatrix} Y_1 - 0.7857Y_2 \\ 1.7857Y_2 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -0.9657 \\ -0.0343 \\ 1 \end{pmatrix}$$

比較のため、前例と抽出位置、日時の同じ標本を選ぶと次のようになる。

	例 1	例 2	例 3	例 4
非 寄 生	0.9815	0.9057	0.7778	0.7639
寄 生	0.0185	0.0943	0.2222	0.2361
標本の大きさ	54	53	54	72

特殊解の部分は、

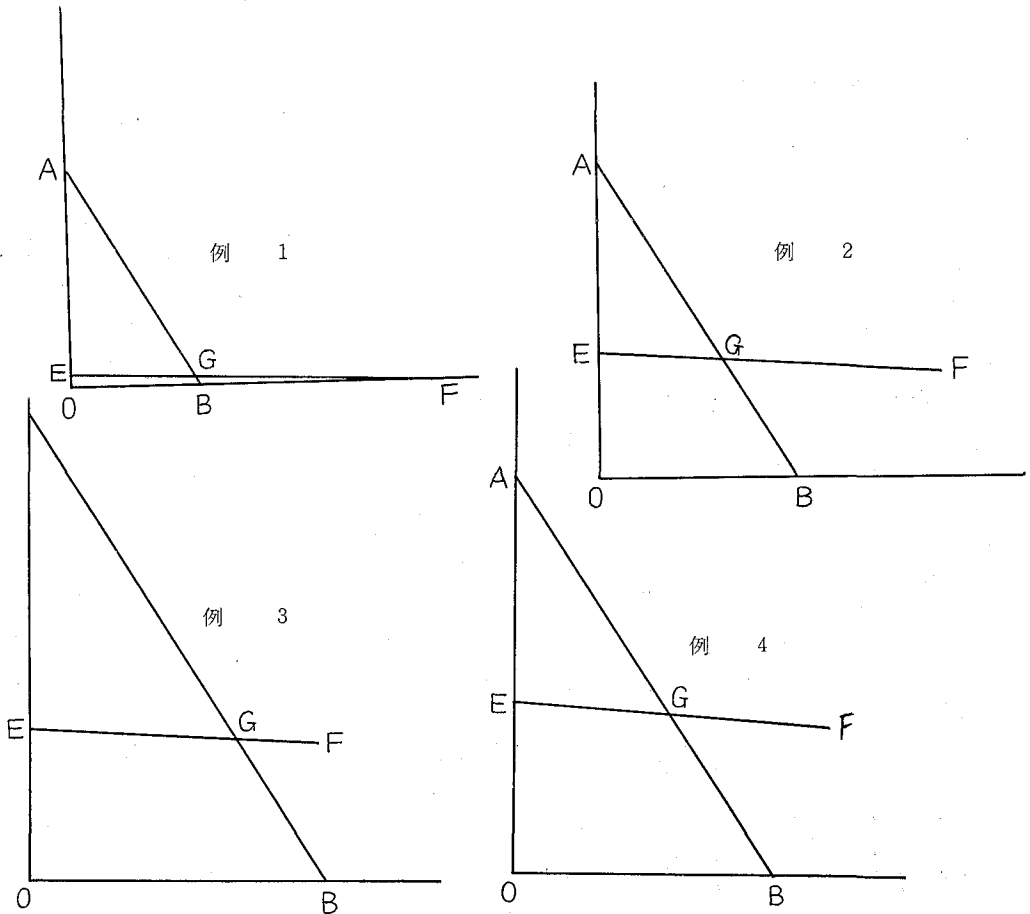
0.9670	0.8316	0.6032	0.5784
0.0330	0.1684	0.3968	0.4216
0	0	0	0

そして $0.970 \geq \lambda_1 \geq 0$ $0.918 \geq \lambda_1 \geq 0$ $0.776 \geq \lambda_1 \geq 0$ $0.758 \geq \lambda_1 \geq 0$

$0.033 \geq p_6 \geq 0$ $0.168 \geq p_6 \geq 0.137$ $0.397 \geq p_6 \geq 0.370$ $0.422 \geq p_6 \geq 0.396$

これから明らかなように、*Trienophorus* という寄生虫学的標識は、アジア系、北米系の混合に関しては、不確定さの程度は非常に大きい。しかし、ウッドリバー群の混合については、殆んど不確定さがないと言っていい程である。これは、言うまでもないことであるが、ウッドリバー群が、この標識では、他の成分個体群よりも、殆んど一様に遠い距離にあるからに外ならない。

このような特殊な情報は、例えば、判別函数のような、一般的により妥当な標識による分析結果の不確定さを減少させる上に有用である。先ず、両者の結果を比較しよう。何れの標識によっても、 p_2 、 p_H のとり得る範囲が規定されている。図で示すと次のようになる。



何れの場合も、判別函数標識は、 (p_2, p_6) 即ち、最初の (λ_1, λ_4) が $\triangle AOB$ の内部にあることを、寄生虫学的標識は、線分 EF 上にあることを要求している。従って、両者の要求が満足されるのは、線分の EG 上である。

次に、直線 EF で $\lambda_1 - \lambda_4$ 平面に直交する平面と、超平面 $-0.389Y_1 + 1.249Y_2 - (0.912\lambda_1 + 0.970\lambda_2 + 0.884\lambda_3 + 0.580\lambda_4) = 0$ (Y_1, Y_2 は判別函数標識の頻度分布) とが、 $\lambda_2 - \lambda_4$ 平面、 $\lambda_3 - \lambda_4$ 平面と交わる点をそれぞれ C, D とすると、 G, C, D, E を頂点とする 4 面体が、北米系各群の混合割合に関する全条件を満足する $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ の値の組を示している。これらの頂点を求めると、

例 1	G	C	D	E	例 2	G	C	D	E
	0.357	0	0	0		0.427	0	0	0
	0	0.329	0	0		0	0.393	0	0
	0	0	0.361	0		0	0	0.431	0
	0.022	0.033	0.033	0.033		0.154	0.168	0.168	0.168
例 3	G	C	D	E	例 4	G	C	D	E
	0.555	0	0	0		0.417	0	0	0
	0	0.510	0	0		0	0.406	0	0
	0	0	0.560	0		0	0	0.421	0
	0.378	0.397	0.397	0.397		0.407	0.422	0.422	0.422

となる。

$f = 1.389Y_1 - 0.249Y_2$; $f = 0.088\lambda_1 + 0.030\lambda_2 + 0.116\lambda_3 + 0.420\lambda_4$ という超平面は、何れの場合も 4 面体を切断せず、従って、頂点の何れかで、最大値、最小値をとる。検討すると、何れの場合も、 D で最大値、 E で最小値をとる。

	例 1	例 2	例 3	例 4
f max	0.056	0.121	0.231	0.226
f min	0.014	0.071	0.166	0.177

そして、アツア系、北米系の混合割合を P_1, P_1' とすると、それぞれ、少なく見積っても、次の値より小さくはないという結論になる。

	例 1	例 2	例 3	例 4
アツア系 $P_1 <$	0.523	0.401	0.044	0.147
北米系 $P_1 <$	0.394	0.549	0.891	0.794

寄生虫学的標識による結果を利用しないときに較べると、不確定さの著しい減少である。勿論、これに抽出変動が附加わるべきことを忘れてはならない。

これらの例では、*Trianaophorus* の寄生が見出された場合を取扱ったが、そうでないとき、 N 尾調査して寄生個体を見出さなかったならば、 $\hat{Y}_2 = \frac{1}{N+1}$ として、この事実を考慮に入れることが出来よう。(統計数値表, 51頁)

便宜のために、Fukuhara の結果を示すと、

	例 1	例 2	例 3	例 4
北米系混合割合	0.318	0.492	0.806	0.668
90% 信頼限界	0.16~0.47	0.29~0.70	0.58~1.00	0.48~0.84

混合割合の推定に際して、成分個体群を定義することが、どのような生態学的意味を持つかは、既に指摘したが、この点から、更に若干の論議をしておこう。

体節の形質にしても、寄生虫学標識にしても、それが各個体に固定するのは、発生初期か、或いは、少なくとも淡水生活期においてである。これは、鱗の或る成長特性を使う場合でも同じである。しかし、現在のところ、成分個体群として定義されるのは、親魚として遡河する群によっている。これが、生態学的意味を持つのは、殆んど 100% の母川回帰を前提しているからである。若し、そうでなければ、たとえ、混合割合

を推定し得ても、世代間関係を追及する基礎にはならないであろう。このような場合、考える局面に入る以前において、定義することが、必要である。寄生虫学的標識の再生産の状況から見るとベニザケに関してはこの帰属性は極めて確実であると考えられる。

親魚として遡河する群で、このように定義された成分個体群には、ベニザケ、シロザケでは、幾つかの年級群が混在している。従って、それは、或る割合での混合を想定しているわけであり、その生態的意味は既に述べた通りであるが、これは、標識頻度に年級変動がないという前提の下に成熟未成熟に拘らず妥当であるが、そうでなければ、遡河直前の成熟群についてだけ妥当であるに過ぎない。

このように、成分個体群の定義には、生態学的問題が密接に絡むわけであるが、それらの生態学的問題の解明は、混合割合の推定が終局的に意図するところである。従って、それは、漸近的に確められ、修正されて行くべきものであろう。

成分個体群の定義に関連して、Fukuhara その他の定義が意味をもつ場合に言及しておこう。混合割合は既述のように、或る混合個体群から任意に個体をとり出したときに、それが或る成分個体群の個体である先験的確率である。従って、若し、成分個体群が相互に混合もしながら分布している全域に亘って、一様な標本抽出が行なわれ、その水域全体としての混合割合を考えるならば、それは、Fukuhara の定義における荷重と一致しよう。逆に言えば、妥当に成分個体群、その標識頻度分布が定義されるならば、そして、全域に亘って積分された混合割合が推定されるならば、成分個体群間の相対的な大きさが推定出来る。この原理は北太平洋オットセイ群（プリピロフ群、コマンドル群、ロベン群）の三陸沖における混合を論ずるときに、合衆国研究者によって使われている*。プリピロフ群は非常に良く研究されており、標識オットセイの日本沖再捕にもとづく混合割合推定に対する傍証として、妊娠率が利用された。アジア系2群については、情報がなく、従って、或る仮説に基づいて不確定ではあるが、妊娠率は100%を越すことがないので、限界値は定まり、それに基づく論議が展開された。その場合もそうであるが、たとえ構造的な不確定さに基づく論議をするにしても、重要なのは、一様な標本抽出が行なわれているか否かということである。一般的に言って、これは、非常に難しいことであり、従って、個体群の相対的な大きさを推定することも容易ではない。生態学的研究としての所謂種族研究においては、他方において、成分個体群の追及ということが、不可欠の要件だから、なお更そうであろう。

要 約

水産資源の調査研究で、形態学的諸形質その他による種族分析と称されている研究は、個体群の間に異質性のあることを指摘するだけに止まっていることが多いが、形質の異質性から、種族或いは個体群の生物学的意味に関する論議を展開するのは、観念的飛躍であるように思われる。考える個体群に生物学的な意味を与えるのは、生活史を通じて数量変動の単位として個体群を追及すること、そのような意味での生態研究全体のあり方であって、標識としての形質は、そのような個体群を追及する手段に過ぎない。そして、このような種族分析が課題として重要性を持つのは、生物学的に極めて近縁でありながら、数量変動の単位としては異なる同種個体群の混合が予想され、かつまた、生物学的に近縁なるが故に、どの形質を標識として採用しても、その頻度分布が個体群の間で重複しているような場合である。この論文は、このような場合の分析手段と方法、調査研究上の問題点を論じている。

標識として用いる形質として、個体群の間でその頻度分布の重複の少ない形質が良いが、混合割合推定量の精度という点から言えば、分布の重複度だけが問題なのではなく、重複の仕方、混合割合自身、その時間的空間的な変化なども関与しており、これらの諸事情を超えて、最良という標識は見出し難い。これよりもより基本的な標識選択の基準は、統計的な意味で確定的な混合割合の推定が可能か否かであり、それが可能であるためには、考える成分個体群における頻度分布から作られる頻度行列の階数が、成分個体群の数より小さくないこと、このために、標識の階級数が成分個体群の数より小さくないことが必要である。更に、統

* 北太平洋オットセイ条約交渉会議・生物学委員会議事録および文書（未公開）

計的に、考える成分個体群の妥当性を検定するためには、標識の階級数が成分個体群の数より大きくなくてはならない。若し、等しいだけならば、考える模型の妥当性が、生物学的に保証されていなくてはならない。

このような条件を満たす標識が採用され、成分個体群から標本抽出していることが確かならば、最尤法によって、標本の提供する情報を完全に利用する推定が可能であり、method of scoring を用いれば、多数の標本であっても、高速計算機による機械的処理が可能である。

採用した標識が、上述の条件を満足しなかったり、或いは、成分個体群からの標本抽出が不確かであれば統計的な意味で確定的な推定が出来ず、程度は事情に応じて様々であるが、推定として構造的な不確かさが残る。この構造的な不確かさは、精度とは別のものであり、たとえ個々の標本の大きさをどんなに大きくしても残るものであり、従って、調査研究としてはより基本的な問題である。普通に用いられる標識には、技術的に、また自然に、その階級数が固定されているものがあり、そのような標識の場合は、幾つかを複合することによってしか、この構造的な不確かさを除き得ない。連続的変量と見做される標識では、その階級分けの仕方に制限はないから、その効用は大きい、階級数を多くすれば、それだけ大きい標本が必要であり、自ら限度があろう。成分個体群からの標本抽出の不確かさに由来する構造的な不確かさを考えると、先ず、調査研究としては、成分個体群追及の努力が払われなければならない。と同時に、構造的な不確かさを常に明確に分析することが重要である。この点は、既往の研究で、一見確定的な推定をすることによって、等閑にされていたところであるが、凸函数の理論を用いて、構造的な不確かさを分析する方法を具体的に、即ち北洋におけるアジアカ・北米系ベニザケの混合問題を素材として、提示している。

調査研究として余儀ない構造的な不確かさをそのままに分析することは種々の前提に基づく一見確定的な結果を得るよりも、遙かに実り多いものである。何故なら、水産資源の調査研究では、調査研究の過程で、資料化されない、或いは定性的な、或いは断片的な情報が可成り得られるのが普通であり、それらを正当に考慮し、より生物学的な洞察が可能になるからである。

最後に、この研究は、北洋資源の調査研究関係者、更には、内海区水産研究所の花岡所長、資源部の人達の不断の支援と協力の賜のものであり、ここに記して、衷心よりの感謝の意を表する次第である。

文 献

参 考 図 書

- 相川広秋：資源生物学，1960。
Rao, C. R: Advanced Statistical Methods in Biometric Research, 1952。
Charles. A, W. W. Cooper and A. Henderson: An Introduction to Linear Programming, 1953。
統計科学研究会編：新編統計数値表，1952。

引 用 文 献

- 1) Bilton, T. H, M. P. Shepard and D. W. Jenkinson, 1958.
The characteristics of the scales of chum salmon taken on the high seas of the North Pacific Ocean in 1956. INPFC* Doc. 218.
- 2) Bilton, T. H. and M. P. Shepard, 1959.
Scale characteristics of chum salmon (*Oncorhynchus keta*) from the North Pacific area in 1957. INPFC Doc. 328.
- 3) Dorn, W. S. 1961.
On Lagrange Multipliers and Inequalities. Operations Research, Vol. 9, No. 1.
- 4) Fukuhara, F. M. and S. Murai, J. Lalanne and S. Sribhidhath, 1959.
Use of morphological differences in determining the area of origin of red salmon taken on the high seas. INPFC Doc. 320.

* International North Pacific Fisheries Commission の略。

- 5) Fukuhara, F. M. 1960.
Analysis of red salmon morphological data—1958. and 1959, Progress Report of Investigation by the United States during 1960. INPFC Doc. 420.
- 6) Hartt, A. C., 1959.
Movement of salmon in the North Pacific Ocean and Bering sea as determined by tagging 1956—1958, INPFC Doc. 327.
- 7) INPFC Secretariat, 1960.
A Preliminary Assessment of the Status of the INPFC Research Program on Questions raised by the Protocol. INPFC Doc. 356.
- 8) Kobayashi, T. and S. Abe, 1958.
On the identification of chum salmon stocks in the North Pacific by means of scales. INPFC Doc. 207.
- 9) Margolis, L., 1958.
The application of parasitological data to the problem of recognizing the continent of origin of stocks of sockeye salmon in the North Pacific Ocean and adjacent waters—a summary of three years' (1955 through 1957) research and addendum. INPFC Doc. 200.
- 10) Margolis, L., 1959.
Report on parasitological studies of sockeye salmon collected in 1958. with some comparison with other years. INPFC Doc. 303.
- 11) Margolis, L., 1960.
Report on the use of parasites to determine the continental origin of sockeye salmon taken on the high seas from May to August, 1959. INPFC. 409.
- 12) Marr, J. C. et al, 1957.
Contributions to the study of subpopulations of fishes; Special Scientific Report, Fisheries No. 208, U. S. Fish and wild life service.
- 12a) Royce, W. F.
Statistical comparison of morphological data.
- 12b) Widrig, T. W, and B. A. Taft.
Measurement of population movement by observation of meristic or morpho-metric characters.
- 13) Ridgway, G. J., G. W. Klontz, Mary LaRoque and E. D. Ullman, 1958.
Serological studies to determine the continental origin of red salmon taken on the high seas. INPFC Doc. 233.
- 14) Ridgway, G. J., G. W. Klontz and C. Matsumoto, 1959.
Intraspecific difference in the serum antigens of red salmon demonstrated by immunochemical methods. INPFC Doc. 313.
- 15) Worlund, D. D. 1957.
Progress Report—The use of multivariate analysis in racial studies. INPFC Doc. 109.
- 16) Worlund, D. D., 1960.
A method for computing the variance of an estimate of the rate of intermingling of two salmon populations. (Mimeograph submitted at the 1960 Annual Meeting)