

分散分析と海岸生態学（2）

大垣俊一

II. 分散分析の前提

前節までに述べてきた分散分析のさまざまな手法を実行するには、いくつか満たすべき前提がある。それらは大きく分けて、分散分析に限らず推測統計一般が要求する条件と、parametrics 型分散分析に固有の条件の 2 種があり、前者には標本の独立性と random sampling、後者には正規分布や等分散性の仮定などが含まれる。

1. 標本の独立性 (independency)

すべての推測統計処理が満たすべき前提として、標本の独立性ということがある。これは、「母集団内のどの標本も、選ばれる確率が等しい」と言い換えられる。つまり、 N 個の標本があれば、それぞれ $1/N$ の確率で選択される、ということである。たとえば海岸のある範囲にコドラートを置いて、範囲内のウニの平均密度を推定しようとするれば、どの地点も等しく選ばれる可能性を持たねばならず、ウニがいそうなところばかりに置いたり、特定の地形のところだけに置くといったことをしてはならない。そして、independency を保証するための方法が、無作為抽出 (random sampling) である。具体的方法としては、たとえば調査範囲を示す地図に格子点を打って番号をつけ、乱数表を用いて必要な数の地点を選び出す。海岸でコドラートを投げるとか、まして、調べやすいところだけに置いて、後でそれがランダムだったとみなす、などというのは random sampling とは言えない。この independency の問題は、統計処理の前提であるから、本来これがどの程度 violate されたらどのような補正が必要になる、といった性質のものではないはずである。たとえば、 t 検定の基準になる t 分布は、正規分布母集団からの random sampling に基づいて描かれているから、random sampling が実現していないと t 分布が使えず、検定不能となる。ただしこの点については、いわゆる interspersion の問題をめぐり、分散分析の歴史の初期に、Fisher と Gossett の論争があった。操作実験においては、処理区とコントロール区がそれぞれ偏らないように、相互散在的 (interspersion) に両者のコドラートを配置することがある。Gossett はこれを支持し、Fisher は random sampling に反するとして反対した。interspersion を前提とした t 分布 (?) は、random sampling を前提とする通常の t 分布とは形が異なる。interspersion の支持者は、相互散在によってすら差が検出されたとすれば、差があることはより確定的である (いわゆる conservative) から問題はないとしている (Hurlbert 1984)。しかし基準分布をケースに応じて変更することに関しては、異論のある研究者も多いことだろう。

ところで案外おざなりにされているが、random sampling を実現するためには、サンプリングのための母集団（あるいは調査範囲）を確定しておかねばならない。そうでないと、どの範囲からランダムに標本を取るか、その作業を始められない。これは単純な密度推定などの場合だけでなく、操作実験によって何らかの処理の効果を調べる場合にも当てはまる。たとえば A 種を導入したことによる B 種の反応といったことも、はじめに設定し、その中にランダムにコドラートを置いたある面積に対する推定値であって、周囲の海岸全体とか、まして A 種と B 種の地理分布全域にわたる値などではない。なお、この点について誤解を招きやすい言葉として、有限母集団 (finite population)、無限母集団 (infinite population) という用語がある。無限母集団とは、標本の抽出の仕方によって、抽出を無限に反復する可能性があるとき、そのように表現するのであって、得られた推定値の適用範囲が無限であることを意味しない。たとえば $10 \times 10\text{m}$ の範囲を $1 \times 1\text{m}$ のメッシュ 100 に区切り、そこから調査区を選べば母集団は 100 で有限である。しかし、位置にこだわらず $10 \times 10\text{m}$ の範囲から $1 \times 1\text{m}$ を選ぶということなら、コドラートの配置の仕方は無限であるから、無限母集団と考えることができる。

2. 正規分布と等分散

parametrics 型の分散分析では、比較する 2 母集団の分散が等しいことが条件となる。また、約 20 以下の小標本の場合、母集団の正規性ということがこれに加わる。その理由としては、前者の場合、2 母集団の分散が顕著に違っていると、大分散の集団からの標本は、小分散の集団がカバーしない '端' の領域から取られるケースが生じやすい。このため、平均値が等しくても「差あり」と認識される確率が高まり、最終判定の際の危険率が不安定化する (大垣 2000 参照)。つまり、実際には差がないのに、差があると誤判断する危険が増すのである。正規分布については、小標本の場合、分散分析の検定操作に用いる t 分布や F 分布は、正規分布母集団からの繰り返しサンプリングに基づいて描かれているから、これを使うためには当然、母集団が正規分布していなければならない。しかし、20~30 以上の大標本を取ることができれば、母集団そのものが非正規分布であっても、平均値の分布は正規分布となり (中心極限定理)、平均値比較は意味を持つことになる。ただし、この正規分布性については、統計の教科書類では、等分散ほど深刻に考えられていないようだ。Sokal & Rohlf (1981) は、一般に標本平均は標本値そのものよりも正規分布に近づく傾向があるので、小標本でも母集団の正規性はあまり問題にならないとし、Underwood (1997) は、正規分布性は小標本の場合も、balanced data (比較する集団の標本数が等しいデータ) ならあまり気にする必要がないので、それを実行するように勧めている。また Zar (1999) は、t 検定は正規分布からのずれに対して頑強 (robust) で、特に標本数相等の両側検定の場合にそうであることは、数多くの研究によって確かめられているとしている。ただし、前 2 著者は、正規分布からの著しいずれは問題があるとして、その対策 (後述) を論じている。

3. 加算性 (additivity)

分散分析の教科書類では、処理の前提のところでは加算性ということもふれられている。これは固定型変数、2元配置以上の分散分析で、繰り返しのない配置の場合に限られる、どちらかといえばマイナーな問題である。今、標本に対し、2つのレベルからなる R の処理 (R1, R2) と、同じく2つのレベルからなる C の処理 (C1, C2) を行って、下記の結果を得たとする。

	C1	C2		C1	C2
R1	5	10	R1	5	10
R2	15	20	R2	15	30

表 8. additive (左) と multiplicable (右)

左表のように、C1 → C2 は+5、R1 → R2 は+10 という、たし算的關係 (additive) で統一されていると、R2-C2 のマスは 20 となり、実際そうになっているから、この場合 R と C の相互作用は、なしとみなされる。しかし、右表のように、R2-C2 のマスが 20 より大きくなっていると、これは R 行の処理と C 列の処理に相互作用があって、互いに強め合って R2-C2 のマスの、20 より大きい値を実現したと考えられる。表では、C1 → C2 は×2、R1 → R2 が×3 と、たし算ではなくかけ算 (multiplicable) の關係になっている。このように相互作用がある場合、ANOVA を型どおりに適用すると、誤差分散の中に相互作用の分散が混入し、両者は区別できない。その結果大きめの V_E で検定することになるので (表 7)、得られる F 値は小さめとなり、検定制度を落とすことになる。繰り返しのある配置の場合は、相互作用を分離することができるので、こうしたことは問題にならない。

4. 前提を満たすための対策

以上に示した分散分析の前提条件が満足されないときは、原則的には、データをとる段階でできるだけ満足するように努めるべきだということになる。特に Underwood (1997) などはこの点を強調している。しかしデータ処理の段階で、前提を満たさないデータに対して検定を有効にするための対策も、種々考えられている。

不等分散については、特に大きな分散が検出された群に対して、異なる種類のデータが混入していないか検討する。また、データの対数変換や平方根変換によって、分散の有意差を消して検定する、という方法が示されている。正規分布からのずれは、標本数を増やしたり、比較する標本集団の数をそろえて影響を少なくしたり、ヤマが分離していそうならば別々の山を別に取り扱う。また、分散と同じく変換の利用や、nonparametrics の使用を勧めているテキストもある。加算性については、各処理群に

対して replicate をとり、相互作用を分離したり、また multiplicable については、対数変換によってデータを additive にしてから検定するという手段も示されている。特に、対数や平方根への変換は、海岸生態学の論文でもよく目にする、一般的な手法である。

Ⅲ. 分散分析と海岸生態学

1. 理論的問題

分散分析は、統計理論の中で最も歴史が古く、かつ最もよく整備された手法であると言われている（粕谷 1998）。しかし、理論の内部に立ち入ってみると、そここに大ざっぱな部分も認められる。たとえば小標本の分散分析の基礎となる t 検定では、t 値の計算のため、本来母集団の分散を用いなければならないところ、標本分散で代用する。もちろん、不遍分散の考え方により、n の代わりに n-1 で割り、理論的に母集団分布に近づくように調整される。しかしそれは、無限に近いサンプルを抜き出したとき、その標本分散と母集団分散の間の関係を正しているということであって、いま現に抽出された 1 標本群の分散を、不遍分散によって補正しても気安めにしか過ぎない。抜き出した 10 回のうち、8 回はまあまあだが、後の 2 回はかなりはずれている、といったことが起こっているはずであり、ここには何らかの危険率が想定されねばならない。また、不等分散検出のための F 検定において、「有意差がなければ同じとみなす」という発想にも問題がある。これについては以前にも論じたことがあるが（大垣 2000）、ここで評価すべきなのは、このような「同じなのにちがうとみなしてしまう危険」（Type I error）ではなく、「ちがうのに同じとみなしてしまう危険」（Type II error）でなければならない。前者でよいのなら、標本数を減らして検出力を下げ、差が見出せなくしてしまえばよいことになる。また、一般に統計のテキストでは、t 検定の前提となる母集団の正規性はあまり問題にならないとしているが、これも何がしかは最終的な仮説棄却の危険率を不安定化させているはずである。これらの問題は、教科書類が指摘するように、一つ一つについては影響は軽微なのかもしれない。たとえば Zar (1999) は、分散不等の影響を示した表において、危険率の変動が 0.07 程度までを「軽微」の範囲として議論している。しかし各要因が積み重なった場合にはどうなるのであろうか。「多重比較」検定の理論において、Duncan 法を否定し、神経質なまでに積み重なりを効果を排除しようとする分散分析が、理論の根本においてこのような大ざっぱなことをやっているとよいのかという疑問なしとしない。この場合、最終的な危険率のゆらぎは、 $P=0.05$ を境に仮説を棄却するということが現実に行われている以上、最大でも 0.045~0.054 の範囲に収まらねばならないと私は考えるが、それは実現しているのだろうか。

「変換」の操作も疑問である。分散不等、非正規性などへの対策として、対数変換や平方根変換が常識のように用いられている。この操作の正当性について、Sokal &

Rohlf (1981) は、われわれが慣れ親しんだ linear ないし算術的スケールのみを用いる科学的根拠はなく、スケールのとり方は任意であってよいと述べる。しかし分散が同じかちがうか、多元配置 ANOVA における相互作用があるかないか、などはそれぞれ一つの自然現象である。それが、変換したら分散差が消えたとか、相互作用なしとみなしてよくなった、というのは、いわば紙の上で自然を変更するに等しい。これは、interporation を行う際、計算の便宜のためにいったん対数軸に変換して回帰式を求めるとか、少数の大きな値によってデータのばらつきが大きくなるので、対数グラフにして収める、といったような、いわば便宜的な変換とは意味が異なる。それらの場合、議論のベースは linear のままである。Underwood (1997) は、分散の差が検出されたら、変換を行う前に分散差の理由を考えよと言う。しかし変換前に分散差があり、変換して差が消えたデータを前にして、我々はどのような議論をすべきであろうか。ここには案外、人間の認識構造にかかわる重大な問題が隠されているのかもしれない。

この場合、分散の差はあるのかないのかと問えば、「差はあるともないとも言える。しかし、同時にあり、かつなしとすることはできない」という答えしかない。たとえば地面を這うアリにとって、目の前の 2 つの小石には明らかな大きさの違いがあるとしても、空を飛ぶトンビの眼には差が認められないかもしれない。この時、アリの視点に立って差を認め、世界像を描くこともできれば、トンビの視点で差なしとし、別の世界像を構築することもできる。両者は異なる見方で自然を見た結果として、同等の価値を持つものである。Sokal & Rohlf (1981) が、linear なスケールのみを正しいとする根拠はないというのは、その意味では正しい。しかしそれらを混在させることはできない。それはたとえば、平面幾何学の体系で問題を解きながら、球面幾何学の定理を持ち込むようなものである。三角形の内角の和は 180° の前提で考えているときに、都合が悪いと 150° にしたり 200° にしたりといったことが許されるはずがない。統計を使用する多くの研究者は、ある場合は linear なまま、またあるときは対数やルートを用い、気ままに座標系を変えながら、目先の技術的な問題がクリアできればそれでよしとしている。そしてそこに、とんでもない認識体系の転換が生じていることに気づいていない。もしこれが大した問題でないと言うのなら、彼らは分散差や相互作用があるのかないのかという問いの意味するものを、深く考えてみるべきではないか。

2. 潮間帯環境への適用

私は以前、ANOVA のような複雑な手法を潮間帯（特に岩礁）に適用することに直感的な違和感があると述べたことがあるが（大垣 1999）、同様の懸念を表明する研究者は他にもある。Shipley (2000) は、野外環境は複雑なので、実験群とコントロール群をランダムに配置しても、大ていはいは何かの要因が両者で分離してしまうと述べている。また Palmer (1988) は、植生調査では検討指標の 'spatial dependence' が見られるのが普通であり、このような状況では、たとえ random sampling を行っても、ANOVA、相関、nonparametrics などを含めた inferential statistics（推測統計）は深

刻な影響を受ける、という重要な指摘をしている。この場合、spatial dependence とは、密度や種組成などが、場所的に偏って分布していることを示す。ここでは、これらの著者の指摘も踏まえながら、環境の複雑さ (heterogeneity) が、分散分析に対してどのような影響を及ぼすかを、海岸生態学の立場から考えてみる。

まず、Palmer (1988) の指摘する 'spatial dependence' は、もちろん潮間帯にも当てはまる。これがどのように検定操作に影響を及ぼすか、Palmer は具体的に論じていないが、私が考えつく一例を示そう。個体群のコロニーにしろ、プールなどの無機環境要因の分布においてもまた、斑状構造 (patchness) は潮間帯に一般的である。たとえば、ヒバリガイモドキのような付着性二枚貝のコロニーが海岸にパッチをなして分布するとき、ここで random sampling によって密度を推定しようとする、パッチに当たったコドラートでは密度が極めて高くなり、はずれたコドラートではゼロに近くなる。その結果、密度の頻度分布を描くと、ピークは、極めて密度の高いところとゼロに近いところに分離した、二山形になるであろう。これは ANOVA の前提となる、正規分布性の violate に他ならない。母集団の正規分布からのずれは大きな問題ではないと言われているが、multimodal など、極端なずれは深刻な影響をもたらすという指摘がある (Sokal & Rohlf, 1981; Underwood, 1997)。そういう場合は、変換や山の分離評価によって対応できるというのだが、前節で指摘したような問題のある変換を、仮に目をつぶって使うとしても、完全に山の分離したデータを正規分布にすることはできない。後者の、分離評価は部分的には対策になりうるだろう。二枚貝コロニーの例で言えば、コロニーの部分と、それ以外を、別々に密度評価するということである。しかし潮間帯における patchness は、目に見えるものばかりではない。砂に潜っている生物の分布や、生物の行動に影響を及ぼす無機環境条件なども、斑状に分布している可能性がある。こういうところにコドラートを置いて、捕食、摂食活動などを調べると、活動値の母集団はパッチ状に分布することになる。Kramer & Schmidhammer (1992) は、 χ^2 検定に対する母集団の spatial dependency の影響を検討し、この検定は、母集団がポアソン分布をしなければ適用不可であるとした。ANOVA の場合に母集団が何分布をしなければならないのか、私には判断できないが、以上述べてきたことからしても、何でもかまわないということではなさそうである。

第二に、これは多元配置の分散分析についてのことだが、生物の反応に影響する要因がいくらかでも考えられる野外環境において、限られた条件だけを取り上げて個々の要因の影響や、その相互作用まで計算することにどれほど意味があるかという疑問がある。time×height×treatment の相互作用などといっても、プールの存否や底質の状態、付着生物の有無など、ほかにいくらかも条件は考えられる。検討していない別の要因が関与しているとすれば、このような検討の枠組みそのものが崩れることになる。実験室内のように、厳密に環境をコントロールできる条件下で、誰が見ても与えられた A 条件と B 条件しか問題にならないことが明らか、というような場合ならば、A×B の相互作用の検討も意味を持つだろう。しかしこのような高次の統計処理を野外に適用することが適切であるかどうか。実際、文献を読んでいても、一応型どおり相互作用も計算してあるが、有効な議論がなされていないという例が多い。

第三に、これは必ずしも分散分析の理論的問題というのではないが、現在の海岸生態学において、ほとんど ANOVA を行う際の前提となっている、コドラートの **dispersion**、**interspersion** の問題に触れておきたい。'dispersion' は、密度やサイズなどを調べるとき、コドラートを固めて置かず、分散させることを指し、'interspersion' は、野外実験において、処理区とコントロール区を相互分散させるように配置することを意味している。これによって、前者では狭い区域の特殊事情に左右されることなく、より広い範囲を客観的に推定することができ、後者では処理区とコントロール区的环境を均一化できると信じられている。しかし再三指摘するように、海岸環境は複雑であって、数多くの環境条件が、いろいろなパターンで入り乱れて存在している。コドラートを固めたためにすべてプールに落ちることもあれば、分散させたためにすべて海藻の上に乗ってしまうこともある。この場合、そのどちらがより頻繁に起こるかということは、一概に言えない。Underwood (1996) は、オーストラリアで巻貝の密度とコドラート離間距離の関係を調べ、50cm 以上離すと個体数の相関は見られなくなり、離れるほど 2 コドラート間の分散が大きくなるという傾向もなかったとしている。一方、Dye (1998) は、南アフリカで、海岸生物種の、種ごとの密度の時間変動を調べ、変動パターンの相関は、地点間距離が増すほど減少すると述べている。私はウニの密度の **variance** が、1 箇所コドラートを接着させた場合のほうが、分散させた場合よりもかえって大きいという結果を得ている。しかしこれらはいずれも、それぞれの地域の別の場所で調べれば、逆の結果が出ることも考えられる。このようなことは、フィールドの性質や、対象とする生物、また採用するパラメーター（密度、サイズ、種組成など）によって変わってくるはずであり、一般化は無理であろう。問題なのは、このような海岸環境の複雑さや、各フィールドの特性といったことを無視して、コドラートが接着していれば即不可とし、分散させてあれば即 OK とする風潮があることである。そしてこの時、批判のために使われるのが、'pseudoreplication' という言葉である。しかし、本当の意味で **pseudoreplication** を避けようと思えば、単なるコドラートの平面分散ではすまず、高さ、湿り気、底質の性質など、あらゆる環境軸を考慮に入れた、多次元空間内の **dispersion**、**interspersion** を実現しなければならない。平面的分散というのは、考えてみると何を分散させているのかよくわからないところがある。

分散分析は、**random sampling**、要因ごとのレプリケート、**balanced data** の実現、コドラートの **dispersion**、**interspersion** など、野外で実際に適用するにはしばしば困難な条件を課し、しかも正規分布の **violate** など、理論的な問題もあって、大方信じられているほど厳密でもない。また、知られている要因に限定した上での仮説検定を前提としているため、要因探索的な研究への使用には限界がある。検定が実行できさえすれば、認識体系をそのつど変更してかまわない、という「変換」の思想も問題である。

これらの難点のうち多くは、潮間帯環境の多様性、異質性 (**heterogeneity**) に起因している。もともと分散分析は、農場技師であった Fisher や Gossett らによって先鞭

をつけられ、またこの手法の、生態学への正確な適用を主張して少なからぬ影響を与えた Hurlbert は湖沼学者である。分散分析は現在、生理学や医学など、実験室系の分野で多用されている。確かに小麦畑や湖や、厳密に条件をコントロールできる実験室内であれば、ANOVA を用いるのは容易であり、適切でもあろう。しかしわれわれのフィールドである海岸でもそのまま通用するかどうかについては、よく考えてみる必要がある。

自分の採用しようとしている方法が、どのような理論的背景を持つかを知ることは、自己の研究に責任を持つことであると同時に、その限界を知って新たなステップを踏み出すための第一歩となる。既成の手法の無批判な踏襲の中からは、オリジナリティのある研究は生まれにくいはずである。

引用、参考文献

- Dye A H 1998. Dynamics of rocky intertidal communities: analyses of long time series from South African shores. *Estuarine, Coastal and Shelf Science* 46: 287-305
- Hurlbert S H 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54: 187-211
- 市原清志 1990. バイオサイエンスの統計学. 南江堂.
- 石居進 1975. 生物統計学入門. 培風館.
- 粕谷英一 1998. 統計の話. 文一総合出版.
- Kramer M & Schmidhammer J 1992. The chi-square statistic in ethology: use and misuse. *Animal Behavior* 44: 833-841
- 大垣俊一 1999. 書評 *Experiments in Ecology*, A. J. Underwood. *Argonauta* 1: 9-14
- 大垣俊一 2000. Parametrics と nonparametrics. *Argonauta* 3: 19-29
- Palmer M 1988. Fractal geometry: a tool for describing spatial patterns of plant communities. *Vegetatio* 75: 91-102
- Shipley B 2000. *Cause and Correlation in Biology*. Cambridge University Press
- Sokal R R & Rohlf F J 1981. *Biometry*, 2nd ed. W H Freeman & Company
- Underwood A J 1997. *Experiments in Ecology*. Cambridge University Press
- Underwood A J & Chapman M G 1996. Scales of spatial patterns of distribution of intertidal invertebrates. *Oecologia* 107: 212-224
- Zar J H 2000. *Biostatistical Analysis*, 4th ed. Prentice Hall