統計グラフと構造指標
――国勢調査統計に取材して――

田口時夫

ローレンツ曲線やジーニ集中度係数は，一般には所得分布の解析や不平等問題の分析等極めて限られた課題に適用される統計手法として理解されている．然しそれ等は氷山の一角にすぎず，その手法に潜む構成論理に気付くならば，その水面下の適用範囲は意外に広いことを知るのである．事実，ジーニ自身による集中度以外の数多くの統計的手法の提案を初めとして，Iyengar, N.S. や Roy, J. による家計分析への適用，Atkinson や Kakwani によるエンゲル弾力性係数や，所得再分配効果の分析は，その経済領域における一歩立ち入った適用といえる．地域分析や社会学及び人口学的分析においても有望性を持つことは，米沢教授や安田教授等によりつとに指摘され，紹介されているところである．

集中解析の方法として一括される上述の諸方法の内部をつらぬく論理構造は，集団構造（田口「経済分析と多次元解析」東洋経済新報社，昭和59年）として測度論的に公理化し，数学的に展開することが出来るが，ここでは一歩進めて更に具体的な諸統計指標を展開することにしよう．これ等の指標はその具体性の故に集団間の構造比較や構造変化の指標として実証的性格を持つものであり，その意味で今日的な課題に沿うものと思われるからである．他方，これ等の指標は極めて日常的一般的な統計グラフの幾何学的特徴と密着した実用的性格を併せ持つものであって，統計の普及や理解の上でも役立つものと思われるからである．

さて，ここで日常的一般的な統計グラフという場合，それは例えば男女の年齢構成を背中合わせのヒストグラムで表したり，男女別に同じ長さの棒状グラフを，年齢階層別の人口比率で仕切ったりするグラフ表現を意味する．又金額分布を例にすれば，輸出，輸入を同一規模の円を用い，それを各商品種別の取引額の全体に占める比率によって扇状に区切る扇形グラフを想定すればよい．これ等のグラフは男女や輸出入間の構造比較を視覚に訴えて行うものであるが，このような機能は男女の年齢順の累積構成比をそれぞ縦軸，横軸でとらえ年齢をパラメータとする曲線表示によっても損なわれない．この時両者の差は，原点と (1, 1) を結ぶ対角線を基準として計測される．これは特にローレンツ曲線における均等線や集中面積の役割と同じものであり，そこに解析的手法の導入への道が拓かれるのである．

パッチ状の母集団からのサンプリング

岸野洋久

通常の場合，自然の動植物はパッチ状に分布している．個体数の動力学を扱うときもこれを考慮に入れる必要が生ずることがある．David と Moore は集合度の尺度として $I=V/m-1$ を提案している．ここで m, V はそれぞれ標本個体数の平均，分散である．その度数分布としては負の二項分布が知られているが，この場合はその級数パラメータを k とすると，$V=m+m^2/k$ となる．従って k と I には $k=m/I$ の関係がある．

一方，パッチ状の分布をパッチの分布 $N_1(dz_1)$（特性測度 $\lambda_1(z_1)dz_1$ のポアソン確率場）と z_1 に中心をもつパッチ内の個体の分布 $N_2(z_1: dz_2)$（特性測度 $\lambda_2(z_2-z_1)dz_2$ のポアソン確率場）でモデル化すると，領域 A 内の個体数 $N(A)$ について

$$\mathrm{E}\,[N(A)]=\int_A \lambda_1 * \lambda_2(z_2)dz_2$$

$$\mathrm{V}\,[N(A)]=\mathrm{E}\,[N(A)]+\int \lambda_1(z_1)dz_1\left\{\int_A \lambda_2(z_2-z_1)dz_2\right\}^2$$

となる．$\lambda_1(z_1)\equiv\lambda_1$, $|A|$ 小，$\lambda_2(y)=\rho\chi_K(y)$ のときは

$$\mathrm{V}\,[N(A)] \fallingdotseq \mathrm{E}\,[N(A)]\,\{1+\mathrm{E}\,[N(A)]\,/\,(\lambda_1\,|\,K\,|)\}$$

である．即ち k は単位面積当りパッチ数とパッチ面積の積になっている．

更にしばしばパッチは階層構造をなしている．z_{j-1} に中心をもつ高位から見て第 $j-1$ 階層のパッチ内の第 j 階層のパッチの分布を $N_j(z_{j-1}:\ dz_j)$（特性測度 $\lambda_j(z_j-z_{j-1})dz_j$ のポアソン確率場）$j=1,\cdots,m$ とすると

$$\mathrm{E}\,[N(A)]=\int_A \lambda_1*\cdots*\lambda_m(z_m)dz_m$$

$$\mathrm{Cov}\,\{N(A),\ N(B)\}=\mathrm{E}\,[N(A\cap B)]+\sum_{j=1}^{m-1}\int\lambda_1*\cdots*\lambda_j(z_j)dz_j\left\{\int_A\lambda_{j+1}*\cdots*\lambda_m(z_m^1-z_j)dz_m^1\right\}\cdot\left\{\int_B\lambda_{j+1}*\cdots*\lambda_m(z_m^2-z_j)dz_m^2\right\}$$

である．和の各項は各階層のパッチからの寄与を表わしている．

この式は系統抽出により個体の分布を見るとき，I にはどの階層あるいはどれ位の規模のパッチが効いてくるかの評価を与えてくれる．また，たとえばライントランセクト法による個体数推定に際し，発見の日変化から調査の分散を求めようとするとき，大きなパッチ構造をしているときは日間の相関も考慮すべきことを教えている．

また最近接距離の分布と同等な $\mathrm{P}\,(N(A)=0)=\phi_m^0\,(0\,;\,A)$ は

$$\phi_1^i\,(z\,;\,A)=\exp\left[-\int_A\lambda_{i+1}(w-z)dw\right],$$

$$\phi_{k+1}^i\,(z\,;\,A)=\exp\left[-\int\lambda_{i+1}(w-z)dw\{1-\phi_k^{i+1}\,(w\,;\,A)\}\right]$$

と漸化的に求まる．

第 k 近接距離の点配置解析への応用

種 村 正 美

平面上の領域 V に散布する N 個体の位置座標が与えられているとする．標本点-個体間最近接距離を X，個体-個体間最近接距離を Y で表し，それらの分布関数をそれぞれ $p(t)$ および $q(t)$ で表すことにする．位置座標の与えられた点配置データにおいて，X が Y よりも一般に有用であることが示されるが，その理由は $p(t)$ および X の密度関数 $f(t)$ の経験分布が精密に計算できるためである．この経験分布はランダム標本点を実質的に無限個にしたことに対応する．

この考え方は最近接距離に限らず，第 $k(\geqq 2)$ 近接距離へも適用できる．そして，それの精密な分布が求まれば，最近接距離のみを利用する場合に較べて，配置パターンに対するもっと詳細な情報が取り出せることになる．

標本点-個体間の第 k 近接距離を X_k で表し，その分布関数を $p_k(t)$ で表すこととする．また X_k の密度関数を $f_k(t)$ とする．$p_k(t)$ は以下のように表すことができる：

$p_k(t)=\mathrm{Pr}\{X_k<t\}=\mathrm{Pr}$ {ランダム点が第 $k-1$ 近接までの個体を除くいずれかの個体と距離 t 内に存在する}

$=\mathrm{Pr}$ {各個体を中心とし，半径が t の円で k 重に覆われる部分にランダム点が含まれる}

$=|\,V\,|^{-1}$（半径が t で各個体を中心とする円板系による k 重被覆領域の面積）

そして，$f_k(t)$ は次のように表せる：

$f_k(t)=dp_k(t)/dt$

$=|\,V\,|^{-1}$（半径が t で各個体を中心とする円板系による k 重被覆領域の周囲長）