メッシュポイントサンプリングの理論とその適用

吉 川 徹* 岡 部 篤 行**

The Theory of Systematic Point Sampling and its Application

Tohru YOSHIKAWA* and Atsuyuki OKABE**

I. はじめに

本稿の目的はメッシュポイントサンプリングの精度について検討することにある。

地域についてのさまざまな研究分野で、緑被率、農地率、土地利用構成比、建ぺい率など、ある地域のなかで特定の属性をもつ領域（ここからは測定対象領域と呼ぶことにする）が占める面積比が問題になる場面は少なくない。しかしこれらの面積比をデジタイザーなどを使って地図から正確に測定するにはかなりの手間がかかる。一方、メッシュポイントサンプリングを利用すればこれらの面積比を比較的たやすく推定できる。特に近年では、さまざまな種類の25mメッシュ、10mメッシュといった詳細なメッシュポイントサンプリングデータが整備されるようになったため、このような既存データが利用できればみずからデータを作成せずに簡単に面積比を推定できる。

たとえば東京都は、都市計画法にもとづいた都市計画基礎調査による土地利用現況図から25mメッシュポイントサンプリングデータを作成し、さまざまな面積比を求めて解析している（東京都都市計画局1988）。このメッシュマップは、縮尺1/2500の用途図、構造・階数図、変化図に25mメッシュのサンプリングポイントを設定し、土地利用、建物用途、建物階数をデータ化したものである。このデータによれば任意の地域（たとえば区、鉄道沿線、用途地域）での土地利用構成比や建ぺい率をたやすく算出できる。したがってこのメッシュマップは都市政策立案のための基礎資料として威力を発揮することが期待できる。

しかし、こうして求めた面積比はあくまで推定値であり、その利用にあたっては精度についての検討がかかせない。また、既存データが利用できず自分でデータを作成する場合には、手間と精度の関係を把握したうえで最適な面積比推定手法を選択しなければならない。そこで本稿では、メッシュポイントサンプリングによる面積比推定の精度についてほかのポイントサンプリングと比較しながら分析する。この目的のため、まず第II章でポイントサンプリングの諸手法について概説して、つづいて第III章ではその精度を理論的に検討する。さらに第IV章では実際の地図についてポイントサンプリングによる面積比推定値の分散を求めて理論的な検討の結果と比較し、最後に第V章で結論をまとめる。

II. ポイントサンプリングの諸手法

メッシュポイントサンプリングはいわゆるポイントサンプリングの一手法である。一般にポイントサンプリングは、地域内に落とした複数のサンプル点のうち測定対象領域に含まれる点の比をもって面積比の推定値とするもので、次の3種類の手法がよく使われる。

* 東京都立大学教養部 * General Education Department, Tokyo Metropolitan University

** 東京大学工学部都市工学科 ** Department of Urban Engineering, Faculty of Engineering, The University of Tokyo

図 1 3種類のポイントサンプリング
(a : 単純ランダム法 b : 層別ランダム法 c : メッシュポイントサンプリング)

<a> 単純ランダム法 (図1 a)

地域内にたがいに独立なサンプル点をランダムに落とす。

 層別ランダム法 (図1 b)

地域を整形区画 (一般に正方形: セルと呼ぶことにする) に分けて, それぞれのセルにひとつずつサンプル点をランダムに落とす。

<c> メッシュポイントサンプリング (システムティック法ともよばれる) (図1 c)

層別ランダム法とおなじく地域をセルに分けたうえで, まずひとつのセルにランダム点をひとつ落とす。さらにセルに対する相対位置がこの点と同じになるように各セルにひとつずつサンプル点を落とす。

III. ポイントサンプリングの精度

つづいてこの章では前章であげた3種類のポイントサンプリングの手法の精度を検討するために, 面積比推定値の期待値と分散を理論的に求めて比較しよう。

まずはじめにこの章で利用する記号を定義する。

n : サンプル点数 (層別ランダム法, メッシュポイントサンプリングではセル数に同じ)

R : 真の面積比

R_i : 第 i 番目のセル内での真の面積比

$\hat{R}_a, \hat{R}_b, \hat{R}_c$: 単純ランダム法, 層別ランダム法, メッシュポイントサンプリングによる面積比推定値

R_{ij} : 第 i 番目と第 j 番目のセルを平行移動で重ねたときに, 重なった測定対象領域がセル全域に占める面積比

$J(U, V)$: 点 (U, V) が測定対象領域にかかれば1, かからなければ0

これらの記号によってポイントサンプリングの手続きは次のように定式化できる。

- ・サンプル点 $(U_1, V_1), \dots, (U_n, V_n)$ を落とす (落とし方によって手法が決まる)。
- ・ $J(U_1, V_1), \dots, J(U_n, V_n)$ を観測する。
- ・ R の推定値 $\hat{R}$ を次式〔1〕によって求める。

$$\hat{R} = \frac{1}{n} \sum_{i=1}^n J(U_i, V_i) \quad [1]$$

この式による面積比推定値の期待値は, 単純ランダム法, 層別ランダム法, メッシュポイントサンプリングとも真の面積比 R に一致する。したがって, $\hat{R}_a, \hat{R}_b, \hat{R}_c$ はすべて不偏推定量であり, この点では3種類の手法に優劣はない。

一方, この3種類の手法による面積比推定値の分散は次のように表わされる (COCHRAN 1953, MARTÈN 1960, KOOP 1976, BELLHOUSE 1981, OKABE 1984)。なお以下では統計量 x の分散を記号 $\text{Var}(x)$ で表わすことにする。

(a) 単純ランダム法

$$\text{Var}(\hat{R}_a) = \frac{R(1-R)}{n} \quad [2]$$

(b) 層別ランダム法

$$\text{Var}(\hat{R}_b) = \sum_{i=1}^n \frac{R_i(1-R_i)}{n^2} \quad [3]$$

(c) メッシュポイントサンプリング

$$\text{Var}(\hat{R}_c) = \frac{R}{n} - R^2 + \sum_{i \neq j} \frac{R_{ij}}{n^2} \quad [4]$$

この分散が小さいほど望ましい手法であると考えられるので、以下ではその大小を比較しよう。

まず式 [3] から層別ランダム法では各セル内での真の面積比によって推定値の分散が影響されることがわかる。分散が最小値（ゼロ）になるのはすべてのセルで R_i が 1 または 0 の場合に限られる。これは、個々のセルをとりだしたときに、全域が測定対象領域内か全域が領域外かのどちらかに限られる場合であり、このときには正確な面積比が求められる。

また式 [4] からメッシュポイントサンプリングではセル相互での測定対象領域の幾何学的関係が分散に影響をおよぼすことがわかる。また層別ランダム法で分散がゼロになるときにはやはり分散は 0 になるが、それ以外にも特殊な場合には分散が 0 になることがある。

さらに、それぞれの方法の分散の間には次の関係が成立する (OKABE 1984)。

$$\text{Var}(\hat{R}_a) \geq \text{Var}(\hat{R}_b) \quad [5]$$

$$\text{Min Var}(\hat{R}_b) = \text{Min Var}(\hat{R}_c) \quad [6]$$

$$\text{Max Var}(\hat{R}_c) = n \text{ Var}(\hat{R}_a) \quad [7]$$

式 [5] に見るように、層別ランダム法の分散は最悪でも単純ランダム法の分散に等しく、一般にはより小さい。しかしメッシュポイントサンプリングの分散は、最小値は層別ランダム法に等しいが、最大値は単純ランダム法の n 倍に達する。この最大値は、すべてのセルで同一の場所が測定対象領域になっている場合、つまり測定対象領域が規則正しく配列していて、その周期にサンプル点間隔が一致する場合に生じる。たとえば建ぺい率推定の場合には、すべてのセルで同一の場所に建築物が建っている場合である。

このようにメッシュポイントサンプリングの精度は測定対象領域の幾何学的な性質に影響されるが、この影響を調べる指標として、層別ランダム法とメッシュポイントサンプリングの分散の差を用いることができる。式 [3] [4] より分散の差 D は次式で表わされる。

$$D = \sum_{i \neq j} \frac{R_i \cdot R_j - R_{ij}}{n^2} \quad [8]$$

この式 [8] の右辺の分子の値をどう評価するかによって、層別ランダム法とメッシュポイントサンプリングの優劣について次の異なった結論が導かれる。

(1) 「優劣はない」

各セル内での測定対象領域の配置がメッシュごとに独立である場合には、両者に差がない。

(2) 「層別ランダム法が優れている」

測定対象領域の境界がセルの境界を横切っている場合には、隣接セル間では独立性の仮定が成り立たない。このことから BELLHOUSE (1981) は、隣接セル間では多くの場合 $E(R_i \cdot R_j) < E(R_{ij})$ であることを示した。

(3) 「建ぺい率推定などの場合にはメッシュポイントサンプリングが優れている」

OKABE (1984) によって建ぺい率推定など測定対象領域がある条件を満たす場合にはメッシュポイントサンプリングが優れていることが証明された。これによると、頂点がグリッド上にある平行四辺形に関し、辺とグリッド軸のなす角度 θ と幅を所与とすると、ある数 n^* が存在して、 $n > n^*$ のときには $\text{Var}(\hat{R}_b) > \text{Var}(\hat{R}_c)$ となる。ただしこれはひとつの平行四辺形に関してである。さらに、 $\pi/4 > \theta > 0.178$ (rad) では n にかかわらず上式が成り立つ。また θ が $\pi/4$ に近いほど $\text{Var}(\hat{R}_c)$ は小さくなる。

表 1 建ぺい率推定での層別ランダム法とメッシュポイントサンプリングの比較

〔岡部他(1984)による〕

(層別ランダム法の分散が小さい件数/逆の件数)

n \ R	0-10%	10-20%	20-30%	30-40%	40-50%
16 ²	2/4	5/5	0/7	3/3	4/1
32 ²	1/5	6/4	5/2	2/4	3/2
64 ²	0/6	4/6	0/7	1/5	1/4
128 ²	0/6	1/9	3/4	2/4	1/4

表 2 用途地域指定面積比推定でのポイントサンプリングの比較

(単純ランダム法が層別ランダム法より分散が小さい件数/逆の件数)

n \ R	0-10%	10-20%	20-30%	30-40%	40%-
20×15	2/12	0/8	0/3	0/1	0/4
40×30	0/14	0/8	0/3	0/1	0/4
80×60	0/14	0/8	0/3	0/1	0/4
160×120	0/14	0/8	0/3	0/1	0/4

(単純ランダム法がメッシュポイントサンプリングより分散が小さい件数/逆の件数)

n \ R	0-10%	10-20%	20-30%	30-40%	40%-
20×15	2/12	1/7	1/2	0/1	0/4
40×30	1/13	0/8	0/3	0/1	0/4
80×60	0/14	0/8	0/3	0/1	0/4
160×120	0/14	0/8	0/3	0/1	0/4

(層別ランダム法がメッシュポイントサンプリングより分散が小さい件数/逆の件数)

n \ R	0-10%	10-20%	20-30%	30-40%	40%-
20×15	6/8	2/6	1/2	1/0	2/2
40×30	7/7	2/6	1/2	1/0	2/2
80×60	9/5	5/3	1/2	1/0	2/2
160×120	9/5	6/2	2/1	1/0	3/1

と(表1), nの増加にしたがってメッシュポイントサンプリングが層別ランダム法よりも優位に立つことがわかる。これは前章の理論的な検討結果と適合している。したがって, 層別ランダム法の分散が単純ランダム法の分散より小さいことを考え合わせると, 建ぺい率推定ではメッシュポイントサンプリングがもっとも分散が小さくなる傾向があることが確かめられる。

(2) 用途地域指定面積比

神奈川県川崎市内の5箇所に設定した2km×1.5kmの長方形の地域内で, 第1種住居専用地域,

IV. 実際の地図を用いたポイントサンプリングの分散の計測

前章で述べたように, メッシュポイントサンプリングによる面積比推定の分散は, 建ぺい率推定などの場合には層別ランダム法よりも小さくなるが, 場合によっては単純ランダム法よりも悪くなる可能性がある。そこでこの章では, 幾何学的性質の異なる2種類の測定対象領域について, 実際に単純ランダム法, 層別ランダム法, メッシュポイントサンプリングによる面積比推定値の分散を求めて, メッシュポイントサンプリングの分散が単純ランダム法の分散を越えることがあるかどうかを調べよう。測定対象としては建ぺい率と用途地域指定面積比を取りあげる。

(1) 建ぺい率(岡部他1984)

東京都と土浦市の1/2,500都市計画基本図に500m×500mの区画を34箇所にわたって設定して, 式〔2〕〔3〕〔4〕によって建ぺい率推定値の分散を計算した。計算結果から層別ランダム法とメッシュポイントサンプリングの優劣を比較する

第2種住居専用地域、住居地域、近隣商業地域、商業地域、準工業地域、工業地域（都市計画道路、鉄道、都市計画公園・緑地を除く）のおのおのが、地域全体に占める面積比を推定対象とした。このため5箇所の地域で4段階のサンプル点数についてそれぞれ40回づつ3種類のポイントサンプリングを繰り返して面積比推定値を求め、その標本分散を計算した。なお、地域によっては指定されていない用途地域があるので、求められた面積比は計30種類である。また、5箇所の地域のすべてに上記の除外される領域があるので、面積比はすべてたがいに独立である。

この実験の結果は表2に示すとおりで、これから次の点を読みとれる。

- ・ サンプル点数が少ない、あるいは面積比がきわめて小さい場合を除けば、単純ランダム法がもっとも分散が大きい。なお、単純ランダム法以外の分散が単純ランダム法を越えたものは、2例を除けば面積比は0.0004未満であり、残りの2例はいずれもサンプル点をもっとも少ない場合に起こっている。
- ・ サンプル点数が少ない場合にはメッシュポイントサンプリングと層別ランダム法の優劣は明確ではないが、きわめて多いときには層別ランダム法が優位である。

これら2種類の測定対象で分散を調べた結果、層別ランダム法とメッシュポイントサンプリングの優劣については反対の結果が得られた。しかし、メッシュポイントサンプリングと単純ランダム法の比較では、両方の測定対象ともメッシュポイントサンプリングの分散は単純ランダム法の分散を下回っている。これは、メッシュポイントサンプリングの分散が大きくなるのは規則正しく並んでいる測定対象領域の周期にサンプル点間隔が一致する場合であり、建ぺい率や用途地域指定面積比など人工的な測定対象領域は一般にその危険が大きいと思われることを考慮すると注目すべきことである。この結果と、OKABE (1984) による、サンプル点が多い場合には建ぺい率だけでなく測定対象領域が平行多角形、円、楕円である場合にもメッシュポイントサンプリングが良好な精度を示すという結果から、ここでとりあげたものの以外の面積比を推定する場合にも、メッシュポイントサンプリングの分散が単純ランダム法を上回る可能性は小さいことが予想される。

V. おわりに

最後にこれまでに述べた内容をまとめ、メッシュポイントサンプリングの精度のめやすについて検討しよう。

メッシュポイントサンプリングの精度は測定対象領域の幾何学的な性質に影響され、理論的には最悪の場合は単純ランダム法よりも精度が悪くなる可能性がある。ただし建ぺい率など測定対象領域が好ましい性質を持っている場合には層別ランダム法にまさる精度が得られ、また一般的な測定対象についても単純ランダム法よりも精度が低下する場合はまれであると考えられる。

このことから、メッシュポイントサンプリングの精度の下限のめやすとして単純ランダム法の精度をと

図2 単純ランダム法で特定の精度を得るために必要なサンプル点数 (n)
(95%信頼度で真の面積比と推定値の差が特定の値におさまるために必要な点数)

ることができると考えられる。単純ランダム法による面積比推定値は二項分布にしたがうので、特定の精度を得るのに必要なサンプル点が計算できる。たとえば、単純ランダム法による面積比推定値と真の面積比との差を95%信頼度で一定限度内におさめるのに必要なサンプル点数を計算すると、図2のようになる。これまでに述べたことから、メッシュポイントサンプリングでもこの図をめやすとすることができると考えられる。たとえば、岡部他(1984)によれば、目でみて建ぺい率を推定した場合には経験を積み約95%の信頼度で8%程度の精度が期待できるとされている。これと同じ精度を単純ランダム法で得るためには図2によれば $n=11^2$ 程度のサンプル点があればよい。メッシュポイントサンプリングでもこの値をめやすとすることができよう。

文 献

- BELLHOUSE, D. R. (1981): Area estimation by point counting techniques. *Biometrics*, **37**, 303-312.
- COCHRAN, W. G. (1953): *Sampling technique*. John Wiley & Sons.
- KOOP, J. C. (1976): Systematic sampling of two-dimensional surfaces and related problems. *Research Triangle Institute*, No. 2500-1123.
- MATÉRN, B. (1960): Spatial variation. *Meddelanden Statens från Skogsforskningsinstitut Stockholm*, **49**, 1-144.
- OKABE, A. (1984): Relative accuracy of systematic point sampling for estimating an area. DP 15, Department of Urban Engineering, The University of Tokyo.
- 岡部篤行・四茂野英彦・金子忠明(1984): 建ぺい率推定の作業量と精度. 日本都市計画学会学術研究論文集, **19**, 265-270.
- 東京都都市計画局(1988): 東京の土地利用—昭和61年東京都区部.

(1990年10月17日受理)