

東京圏における職業別従事者割合と世帯規模別割合の相関係数

小西 純

Correlation coefficient between workers ratio by occupation and ratio by household size in the Tokyo area

Jun Konishi

Abstract: The purpose of this paper is to conduct a case study on fluctuation of correlation coefficient in the Tokyo area using grid square statistics. The correlation coefficient in the Tokyo area between the ratio of the number of manufacturing process workers and the ratio of the number of single-person households on the results of the 2015 population census by municipality is -0.68. On the other hand, the correlation coefficient is -0.33 on the results by Cho and Aza. In this paper, we calculate the correlation coefficient of arbitrary area by moving window method using grid square statistics on the 2015 population census and GIS, and consider the spatial scale about correlation coefficient.

Keywords: 相関係数 (correlation coefficient), 可変単位地区問題 (modifiable area unit problem), 移動窓法 (moving window method), 地域メッシュ統計 (grid square statistics)

1. はじめに

本稿の目的は、地域メッシュ統計データを用いて、東京圏における相関係数の変動についてケーススタディを行うことである。

都道府県や市区町村などの地域別に集計された統計データを利用した分析については、集計された地域単位のサイズで分析結果が異なる「可変単位地区問題」に注意する必要がある。小西 (2019)は、平成 22 年国勢調査における既存集計地域区分の市区町村別集計、町丁・字等別集計、地域メッシュ統計結果を利用して東京圏（埼玉県、千葉県、東京都（島嶼部を除く）、神奈川県）における世帯人員別世帯数の割合と住宅の延べ面積別世帯数の割合の相関係数を計算し、概ね「集計地域区分が大きいほど相関係数が高い」という傾向があることを確認している。例えば、世帯人員

1 人の世帯数の割合と延べ面積 0~29 m²の住宅に住む世帯数の割合の相関係数は、市区町村別 (0.950) > 町丁・字等別 (0.853), 地域メッシュ別 (0.766) である。さらに、この傾向を踏まえて、世帯人員 1 人の世帯数と延べ面積 20~29 m²の住宅に住む世帯数について、地域メッシュ統計データを利用して任意の地域単位で集計後割合を算出し、東京圏における相関係数を計算して、集計地域単位が大きくなると平均の水準は一定となり、ばらつきが大きくなる傾向があることを明らかにしている。

本稿では、小西 (2019)と同様の分析手法を用いて職業別従事者と世帯人員別世帯数の割合の相関係数を確認し、空間スケールについて考察する。

2. 分析方法

2.1 可変単位地区問題

地域別に集計された統計データを利用して分

小西 純

(公財)統計情報研究開発センター

E-mail: junkonishi@sinfonica.or.jp

析する際の地域単位を本稿では「分析地域単位」とする。地理情報を用いた分布図の作成や、各種の空間データ解析を行う際に、分析地域単位の違いが分析結果に違いをもたらすことがある。これらの分析単位によって結果が変わる問題は空間スケール問題と言われ、可変単位地区問題 (MAUP : Modifiable Area Unit Problem) のひとつとして知られている (Openshaw (1984), 中谷 (2015))。

2.2 移動窓法

平成 27 年国勢調査の 2 分の 1 地域メッシュ別の統計データを利用して、任意の分析地域単位による相関係数を計算する。任意の分析地域単位の集計に当たっては、移動窓法という空間分析手法を用いる。

移動窓法とは、部分地域における統計量と捉えることもできる。濱崎・玉川(2005)は同様の方法を用いて、道路幅員の空間分布について集計後コンターを作成し、道路幅員と土地利用変化の関係に関する分析を行っている。

移動窓は分析地域全域の一部を覆う領域を表す。この領域(「窓」)はある位置から他の位置に移動し、窓には円形か四角形が使われることが多い。多くのアプリケーションソフトウェアでは、窓は分析範囲全体を規則正しいステップで動き、各場所で処理(例えば窓内にある事象の平均の計算など)を実行する (Lloyd(2010))。

移動窓法による集計値の計算においては、下準備として地域メッシュ区画単位で行うのではなく、各地域メッシュ区画の図形重心点を作成し、統計値を持たせた。統計値を属性値として持つ重心点を中心とする半径 Xm の円の範囲(窓)を分析地域単位として設定し、その範囲内にある重心点の統計データを集計した。図-1 に示すように分析地域単位を設け、さらにその範囲を固定させるのではなく、移動させ重ねながら統計データを集計している。以下、この移動窓の半径を「バンド幅」という。

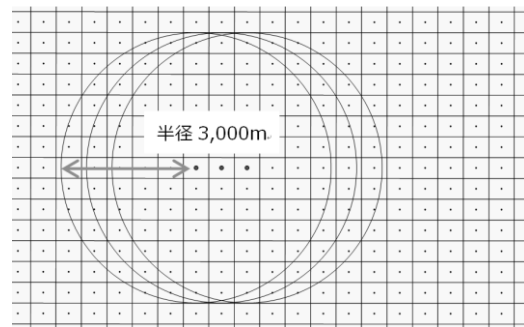


図-1 移動窓法 (バンド幅 3000m)

2.3 ランダム抽出点

移動窓法による集計を行うと図-1 のように隣接する分析地域単位ではほとんどの重心点が重複するため、全ての重心点を用いて相関係数を計算するとその値は高くなる。本稿ではこの影響を取り除いて東京圏における任意の分析地域単位別の相関係数を計算するため、分析地域単位のサイズごとに空間的に互いに独立でランダムな点を作成した(以下、「ランダム抽出点」という)。

地域データは集計する位置によってデータ値の傾向が異なるため、多くのパターンについて東京圏における相関係数を計算することとし、ランダム抽出点はバンド幅毎に 100 組作成した。

表-1 は東京圏におけるバンド幅別のランダム抽出点の平均である。図-2 はバンド幅 3000m のランダム抽出点の例である。

表-1 バンド幅別ランダム抽出点数の平均 (東京圏)

バンド幅	750 m	1500m	3000m	4500m	6000m	7500m
抽出点数平均	3035.7	916.8	263.1	132.7	79.4	52.3

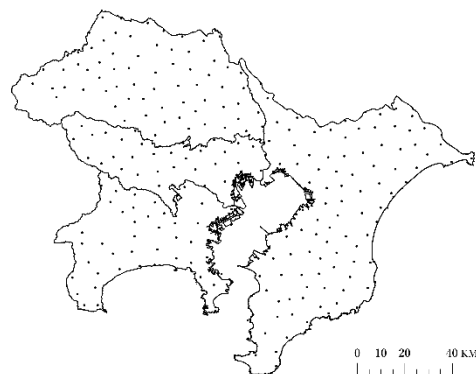


図-2 ランダム抽出点の例 (3000m)

3. 適用例

3.1 既存集計地域区分別の分析

表-2 は東京圏における市区町村別の職業別従事者割合、世帯人員別世帯数割合の相関係数を計算した結果である。相関係数が ± 0.5 以上のセルにはハッチをかけている。世帯人員1人の世帯は管理的職業、専門的・技術的職業、事務従事者などいわゆるホワイトカラーの職業従事者と正の相関関係があることを示しており、生産工程、輸送・機械運転、建設・採掘、運輸・清掃・包装等従事者などのブルーカラーと負の相関関係を示している。またこれらの職業従事者は2人世帯、3人世帯、4人以上世帯と正の相関関係を示している。

表-2 東京圏における市区町村別世帯人員別世帯数の割合と職業別従事者の割合の相関係数 (N=242¹)

	世帯人員 1人	世帯人員 2人	世帯人員 3人	世帯人員 4人以上
管理的職業	0.54	-0.35	-0.53	-0.55
専門的・技術的職業	0.58	-0.47	-0.42	-0.43
事務	0.53	-0.49	-0.32	-0.33
販売	0.38	-0.31	-0.17	-0.21
サービス職業	-0.07	0.33	-0.10	-0.17
保安職業	-0.14	0.21	0.14	0.07
農林漁業	-0.46	0.34	0.26	0.33
生産工程	-0.68	0.48	0.57	0.59
輸送・機械運転	-0.68	0.55	0.56	0.55
建設・採掘	-0.66	0.57	0.52	0.52
運搬・清掃・包装等	-0.62	0.52	0.53	0.50

表-3 は東京圏において町丁・字等別に割合を算出して相関係数を計算した結果であり、ほとんどの項目において市区町村別よりも低い相関係数を示し、相関係数が ± 0.5 以上となる項目は存在しない。

このように集計地域単位によって職業別従事者割合と世帯人員別世帯数割合の相関係数の水準は異なる。

表-3 東京圏における町丁・字等別世帯人員別世帯数の割合と職業別従事者の割合の相関係数 (N=21433²)

	世帯人員 1人	世帯人員 2人	世帯人員 3人	世帯人員 4人以上
管理的職業	0.11	0.00	-0.09	-0.14
専門的・技術的職業	0.30	-0.17	-0.15	-0.30
事務	0.14	-0.08	-0.02	-0.17
販売	0.12	-0.02	-0.03	-0.18
サービス職業	0.00	0.17	-0.07	-0.09
保安職業	0.01	-0.04	-0.01	0.02
農林漁業	-0.32	0.08	0.18	0.39
生産工程	-0.33	0.16	0.21	0.31
輸送・機械運転	-0.26	0.16	0.16	0.22
建設・採掘	-0.25	0.10	0.16	0.25
運搬・清掃・包装等	-0.25	0.24	0.13	0.16

3.2 任意の分析地域単位別の相関係数

前節では集計地域区分が大きいほど相関係数が高くなることを確認したが、市区町村や町丁・字等のサイズは地域によって異なり、境界の区切り方によって相関係数が異なることもあると考えられる。

このため本節では任意の分析地域単位における相関係数について分析する。分析は移動窓法により、そのバンド幅は750m, 1500 m, 3000 m, 4500 m, 6000 m, 7500 mの6種類とした。

隣接する重心点における割合はほぼ同じになるため、集計地域単位が独立となるランダム抽出点と重なる重心点に収録した2変数の割合を東京圏においてリスト化し、相関係数を計算した。

ランダム抽出点1組ごとに2分の1地域メッシュとバンド幅6種類の計7種類の分析地域単位について東京圏における相関係数を計算した。

統計項目は市区町村別の相関係数が高い、世帯人員1人世帯、世帯人員4人以上世帯、専門的・技術的職業従事者、生産工程従事者とした。

丁・字等を除く。

¹ 島嶼部を除く

² 島嶼部、秘匿地域、一般世帯数0、15歳以上就業者数0の町

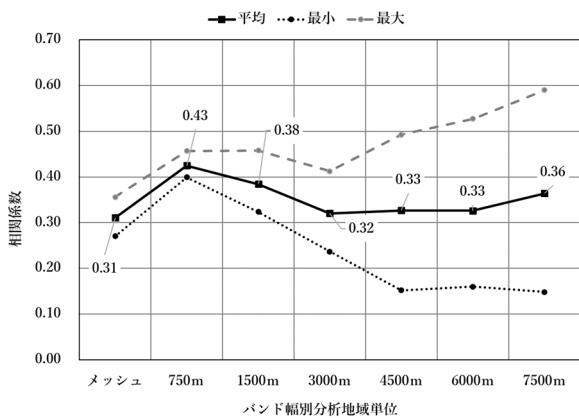


図-3 バンド幅別世帯人員1人世帯数割合と専門的・技術的職業従事者割合の相関係数

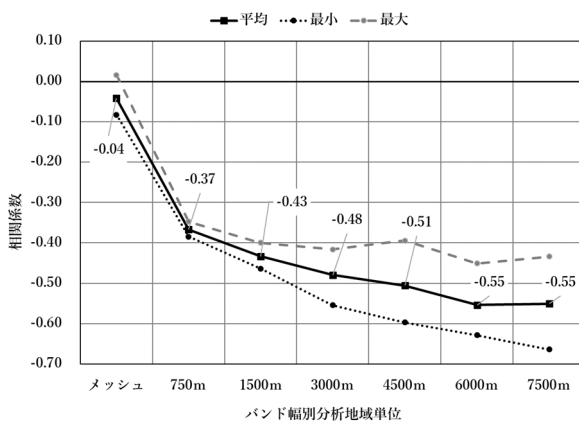


図-4 バンド幅別世帯人員1人世帯数割合と生産工程従事者割合の相関係数

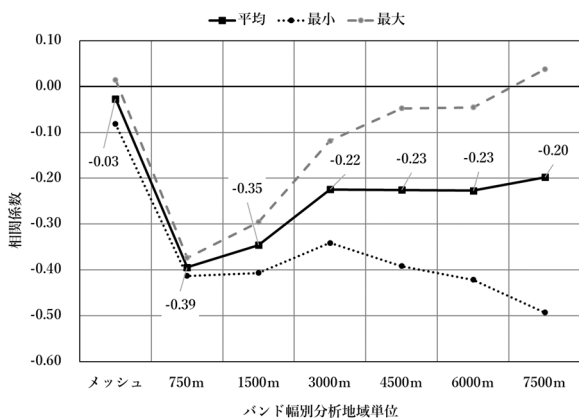


図-5 バンド幅別世帯人員4人以上世帯数割合と専門的・技術的職業従事者割合の相関係数

世帯人員1人世帯数の割合は、専門的・技術的職業従事者と正の弱い相関関係、生産工程従事者の割合と負の相関関係がある。(図-3, 図-4)。4人

以上世帯数の割合は1人世帯数とは逆の傾向を示し(図-5, 図-6), 専門的・技術的職業従事者とは負の弱い、生産工程従事者とは正の相関関係を示す。これらのグラフに共通して「相関係数の平均は分析地域単位が大きくなると一定の水準となり、最小と最大の差は大きくなる」という傾向が見られる。

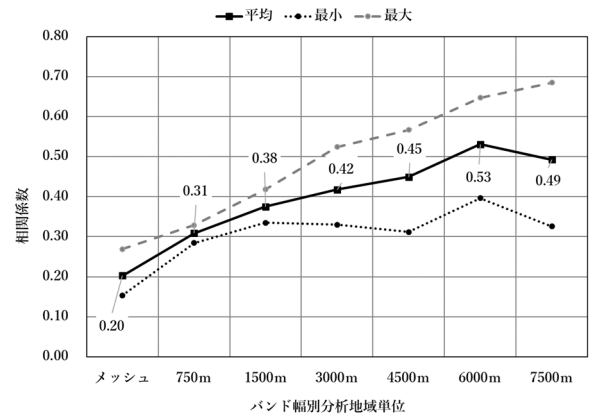


図-6 バンド幅別世帯人員4人以上世帯数割合と生産工程従事者割合の相関係数

4. おわりに

職業別従事者数の割合と世帯規模別世帯数の割合においても平均はある程度の水準で一定となり最大、最小の差などばらつきが大きくなる傾向があることが確認できた。このことから集計地域の「位置」により相関係数の動向は異なり、これは統計項目の空間的分布の偏在性の影響が大きいと推察できる。

参考文献

- 小西純,2019.「東京圏における年齢別単身世帯と住宅の延べ面積の地域別相関係数の地図化」『日本統計学会誌』49-1, 1-32. (印刷中)
- 中谷友樹,2015.「空間分析におけるスケール」『地理情報科学 GIS スタンダード』浅見泰司, 矢野佳司, 貞広幸雄, 湯田ミノリ編, 古今書院
- 濱崎直樹,玉川英則,2000.「ポテンシャルコンターを用いた道路と土地利用変化の関連性の分析」『GIS—理論と応用』8-2, 61-67.
- Christopher D. Lloyd,2010. *Spatial Data Analysis: an Introduction for GIS Users*, Oxford University Press.
- Stan Openshaw ,1984. The Modifiable Areal Unit Problem, *CATMOG*,38, GeoBooks.