

# Small area estimation based on linear mixed model with spatially correlated two-fold random area effects

東京大・経済・院 川久保 友超

東京大・経済・院 菅澤 翔之助

東京工業大・社会理工 小笠原 浩太

## 小地域推定とは？

サンプルサイズが小さい地域（小地域）の何らかの母数（平均など）を推定する際、標本平均のようなサーベイデータによる direct estimator を用いると、推定量の分散が大きくなり推定が不安定になる。そこで、地域効果を変量効果（random effect）として組み込んだ混合効果モデルを仮定し、これにもとづいた model based estimator を小地域母数の推定量として用いることで、推定のリスクを改善することができる。

以下の Fay–Herriot モデルは、地域レベルのデータのみが利用可能な場合の小地域推定モデルとして、最も基本的である。

$$y_i = \theta_i + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, d_i)$$
$$\theta_i = \mathbf{x}_i^t \boldsymbol{\beta} + v_i, \quad v_i \sim \mathcal{N}(0, \tau^2), \quad i = 1, \dots, m$$

興味の対象は小地域母数  $\theta_i$  の推定で、 $y_i$  はその direct estimator、 $\varepsilon_i$  は分散  $d_i$  が既知の標本誤差である。 $\theta_i$  には linking model として上式 2 行目のような構造を入れるが、ここで  $\mathbf{x}_i$  は地域  $i$  における補助情報、 $\boldsymbol{\beta}$  は未知の回帰係数、 $v_i$  は地域効果で分散  $\tau^2$  が未知の変量効果を仮定する。このモデルから得られる  $\theta_i$  の推定量（予測量）を、小地域母数の model based estimator として利用すると、これは direct estimator  $y_i$  よりも推定リスクが小さくなる。

## 空間相関を考慮に入れた小地域推定

古典的な小地域推定においては、データは地域間で互いに独立であると仮定している。上述の Fay–Herriot モデルにおいては、 $\varepsilon_1, \dots, \varepsilon_m, v_1, \dots, v_m$  は互いに独立である。しかし、データに空間的な相関が見られたり、データの構造が地域によって異なったりする場合、空間情報を無視したモデリングは不適切である。そこで、空間情報を組み込んだ小地域推定のモデリングとして、変量効果に simultaneously autoregressive (SAR) process と呼ばれる相関構造を仮定する方法を考える。

本研究では、地域（area）がさらに小さな sub-area に分かれている場合を想定し、地域効果と sub-area 効果を二重に入れた以下の小地域推定モデルを組み、それぞれの変量効果には SAR の相関構造を仮定する。

$$y_{ij} = \theta_{ij} + \varepsilon_{ij},$$
$$\theta_{ij} = \mathbf{x}_{ij}^t \boldsymbol{\beta} + v_{1i} + v_{2ij}, \quad i = 1, \dots, m; \quad j = 1, \dots, n_i,$$

$i$  は地域をあらわす添字、 $j$  は sub-area をあらわす添字、 $v_{1i}$  は地域  $i$  の地域効果、 $v_{2ij}$  は地域  $i$  の  $j$  番目の sub-area 効果をあらわす。このモデルにもとづいた小地域母数の推定量（予測量）を model based estimator として提案し、その推定リスクを評価する。さらに実データ例として、1930 年の東京市における死亡データへの応用を行う。