

東京農業大学 正員 鈴木 忠義
 東京工業大学 正員 肥田野 登
 (株)日通総合研究所 正員 ○長谷川 隆

1. はじめに

従来の地域計画では1つの手法として各種の計量分析が行われた。しかし、これらの分析の限界に関しては、分析過程全般に渡って存在する様々な不確実性のために明確であるとは言えず、分析結果の有効性について判断することは難しい。これらの疑問に答えるために、データの想定誤差に基づいた各モデル自体の精度分析が行われてきたが、インプット情報としての既存統計データの精度に関する分析はまだ不十分であると言える。

本研究では、以上の背景を踏まえて人口関連データの誤差の傾向を定量的に把握し、地域計量分析で必要となるデータ誤差の水準を導出してこれらについて分析の限界を明確にするために1つの試みを行った。ただし、人口関連データとは、国勢調査、住民基本台帳人口調査、事業所統計調査、工業統計調査および商業統計調査による人口、従業者数を示すものとする。

2. 誤差分析の理論

1) 誤差の定義

ここでは調査の真の対象、時点と該当調査のそれとの乖離のために生じる真値と実測値との差を誤差の中に含めた広義の誤差を扱う。定義式を以下に示す。

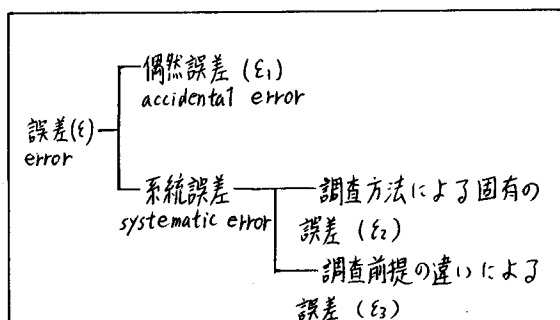
$$X = X^* + \varepsilon \quad (X: \text{実測値}, X^*: \text{真値}, \varepsilon: \text{誤差})$$

2) 誤差の分類

本研究では誤差を以下のように分類した。

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \quad (\text{く図-1を参照})$$

- ε_1 : ランダム誤差
- ε_2 : 調査方法にのみ依存
- ε_3 : 調査方法、地域の特性の両方に依存



く図-1> 誤差の分類

3) 分析の仮定

ここでは1つの調査を1つの単位と考え、その際に発生する相対誤差 λ の値を1つの標本と考えた。定義式を以下に示す。

$$\lambda_j = \varepsilon_j / X^* \quad (j = 1, 2, 3)$$

$$\lambda = \lambda_1 + \lambda_2 + \lambda_3$$

分析の仮定はく図-2>で示す。

4) 相対誤差の分散、期待値の推計法

本研究ではほぼ同一対象に対してほぼ同一時点で調査を行ったと考えられる2つの調査結果の乖離率以下の乖離率の値としてと見え、それらの値を各地域特性ごとに考察することにより相対誤差の分散、期待値を推計した。(計算の過程はここでは省略する。)

(仮定1)

	期待値	分散
λ_1	0	σ^2
λ_2	θ_{λ}	—
λ_3	Δ_{λ}	—

λ : 調査方法
 i : 地域類型
 ($i=1, 2, \dots, I$)

(仮定2)

$$\text{Cov}(\lambda_{1i}, \lambda_{1j}) = 0 \quad (i \neq j)$$

(仮定3)

$$E(\lambda) = \theta_{\lambda} \quad (\text{全サンプルによる期待値})$$

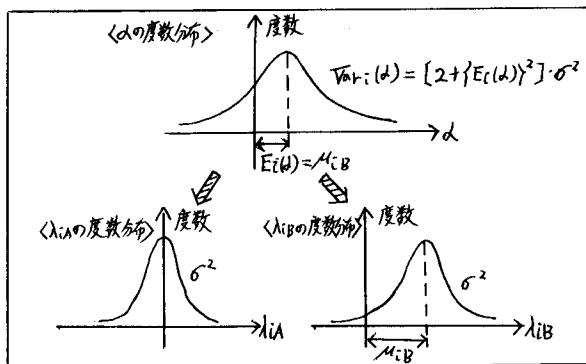
く図-2> 分析の仮定

$$\text{乖離率 } d = (\lambda_B - \lambda_A) / \lambda_A$$

(λ_A : A調査(基準調査)の実測値
 λ_B : B調査の実測値)

ここで、 $\theta_A = \Delta_{iA} = 0$ ($i=1,2,\dots,I$)と仮定すると、次式より相対誤差の分散 σ^2 と期待値 μ_{iB} の値が求まる。〈図-3〉はその概念図を示す。

$$\begin{cases} \mu_{iB} = \theta_B + \Delta_{iB} = E_i(d) \\ \sigma^2 = \text{Var}_i(d) / (2 + \mu_{iB}^2) \end{cases} \quad (i=1,2,\dots,I)$$



〈図-3〉 乖離率の分散と相対誤差の分散との関係

3. 誤差の大きさの定量的把握

ここで用いた地域類型は、地域特性を表わす8個のマクロ指標を用いて主成分分析を行うことにより導出した。

国勢調査を基準とした各乖離率を分析することにより、以下に示す結果が得られた。

1). 製造業、卸売小売業の従業者数は、事統の方が工統、商統に比較して調査法固有の系統誤差が大き

2). 人口、卸売小売業従業者数では、過疎地域で系統誤差が大きく、製造業では、都市業務地域で系統誤差が大き

3). 人口は、市区町村レベルでも信頼できる。

4). 一般に各従業者数とも市レベルから区町村レベルにアグリゲート

の水準を下げると精度が急激に悪化し、ここに統計データとしての1つの限界が存在すると考えられる。

5). 項目間のアグリゲートを行った場合には、産業大分類から3小以下の水準にかけて、上記と同様の理由でデータとしての限界が存在すると考えられる。

6). 時系列的に統計データの精度は向上しており、特に市区町村レベルでの精度の向上が顕著である。

4. 計量経済モデル分析における人口関連データの必要精度およびまとめ

本研究では、上越地方(新潟県)を対象とした計量経済モデルに上で求めたデータ誤差を考慮することにより、人口関連データに関して相対誤差の標準偏差の値で8%前後以下の分析の限界であり、これは使用データとして、従業者数において2次、3次産業全体のアグリゲート水準、また市、郡レベルの水準以上でなければならぬ(1)事加示されたわけである。(詳細は当日に発表の予定)

今後の課題としては、人口関連データ以外の主要データに関して、この誤差の水準、性質を把握することが必要とされた。(参考文献)久保田、他「データ誤差の存在を仮定した地域整備効果分析手法に関する研究」第30回年次学術、或

		系 統 誤 差				偶然誤差
		都市業務型 ΔI	都市住宅型 ΔII	地方中核型 ΔIII	過疎農村型 ΔIV	σ
住民基本台帳人口		$\Delta 0.6$ (%)	$\Delta 0.5$ (%)	0.2 (%)	2.0 (%)	1.1 (%)
事業所統計 従業者数	鉱業	$\Delta 36.7$	$\Delta 33.5$	40.9	31.1	24.6
	建設業	$\Delta 4.6$	$\Delta 7.8$	8.1	5.2	$\Delta 12.8$
	製造業	$\Delta 4.2$	0.5	2.3	1.7	$\Delta 5.9$
	卸売小売業	0.8	$\Delta 2.7$	$\Delta 2.4$	3.2	3.6
	金融・保険業	20.2	8.4	$\Delta 20.0$	$\Delta 5.6$	$\Delta 12.2$
	不動産業	$\Delta 16.8$	10.7	$\Delta 3.8$	8.1	46.1
	運輸・通信業	1.7	$\Delta 5.4$	$\Delta 4.8$	5.2	$\Delta 11.7$
	観光・水道	$\Delta 5.5$	$\Delta 3.6$	7.6	$\Delta 0.2$	$\Delta 4.6$
	サービス業	$\Delta 22.4$	$\Delta 4.6$	$\Delta 3.3$	9.0	$\Delta 1.9$
	公務	$\Delta 7.1$	$\Delta 7.2$	8.2	6.5	$\Delta 17.2$
工業統計 従業者数		$\Delta 12.3$	$\Delta 0.7$	7.9	6.3	$\Delta 16.0$
商業統計 従業者数		$\Delta 4.4$	$\Delta 4.9$	0.9	6.5	$\Delta 7.1$

〈表-1〉 地域特性と誤差(国勢調査の結果を基準、昭和50年) 市区町村レベル