

副標本による標本誤差の計測

山口幸三*

要旨

公的統計を作成する一部の統計調査では、副標本によって標本誤差を計測している。副標本による方法では、複雑な標本設計をしている場合などでも、容易に標本誤差を計測できる長所がある。一方で、副標本は、それぞれ標本構造を同質にするため、副標本の組数を多く設定することができず、標本誤差がばらつくという短所がある。

本稿では、大規模標本をもつ数年間隔で周期的に実施される統計調査において、副標本による標本誤差がばらつくのか、そのばらつきはどの程度なのかを検証した。検証では、社会生活基本調査を用い、いくつかの異なる副標本に分ける方法によるシミュレーションを行った。その結果、副標本による標本誤差がばらつくこと、副標本にするには、副標本の組数だけでなく、大きさも考慮する必要があることを確認した。併せて、安定的な標本誤差を求める方法について提示した。

キーワード

社会生活基本調査, 標本誤差, 副標本, ブートストラップ法

1. はじめに

公的統計を作成する統計調査のうち、社会生活基本調査、就業構造基本調査、労働力調査などは、副標本によって標本誤差を計測している。副標本による標本誤差¹⁾は、複雑な標本設計をしている場合や集計項目が多い場合でも、容易に求めることができる長所がある。一方で、各副標本は、それぞれの標本の値から母集団の値を推定ができるように、標本構造を同質になるようにしているため、副標本の組数を多く設定することができない。副標本の組数が少ない場合には²⁾、標本誤差の推定値がばらつくという短所がある。毎月調査している労働力調査については、月々の

標本の大きさは変化しないので、月々の標本誤差も大きくは変化しないはずである。しかしながら、8組の副標本による標本誤差の推定値は、月々大きくばらついていることがわかっている（古橋・岩永，1991³⁾）。

標本調査においては、調査結果の精度を標本誤差の推定値によって示している。その標本誤差が、副標本の大きさや構成の違いによってばらつくとするならば、そのばらつきの有無や大きさを確認し、安定的な標本誤差を求めることが必要と考える。毎月調査される標本調査では、月々の標本誤差によってばらつくことを確認することができるものの、数年に1回しか調査されない標本調査では、そうしたことはできない。

そこで、大規模標本をもつ数年間隔で周期的に調査される統計調査において、事後的に設定する副標本による標本誤差がばらつくの

* 正会員，総務省統計研修所
〒185-0024 東京都西国分寺市泉町2-11-16
e-mail : k3.yamaguchi@soumu.go.jp

か、ばらつきの大きさはどの程度なのかを、社会生活基本調査を用い、シミュレーションによって検証する。併せて、いくつかの異なる副標本の構成を試し、標本誤差のばらつきがどのように異なるのかも検証する。その上で、安定的な標本誤差を求める方法について考察する。

2. 社会生活基本調査の標本誤差

シミュレーションによる検証においては、平成13年社会生活基本調査の調査結果を利用する。検証方法について論ずる前に、その前提となる平成13年社会生活基本調査の標本設計、調査結果の推定方法、標準誤差の推定方法について簡潔に説明しておく。

2.1 標本設計

平成13年社会生活基本調査(以下「社会調」という。)では、第1次抽出単位を平成7年国勢調査調査区(以下「調査区」という。)とし、第2次抽出単位を世帯とする層化2段抽出法によって標本抽出している。第1次抽出では、都道府県ごとに確率比例抽出により調査区を抽出し、第2次抽出では、等確率無作為抽出により、各調査区から12世帯を抽出している。

調査対象は、抽出された世帯にふだん住んでいる10歳以上の世帯員全員である。

(1) 調査区の抽出(第1次抽出)

都道府県ごとに、そこに含まれる調査区⁴⁾を次の基準により配列し、この配列を基に、各調査区の人口を累積し、累積した人口に基づく確率比例系統抽出により、調査区を抽出している。

- ① 大都市圏に含まれるか否か
- ② 市町村の人口階級
- ③ 調査区の特長(国勢調査調査区番号の後置番号)
- ④ 市区町村コード
- ⑤ 国勢調査調査区番号

(2) 調査世帯の抽出(第2次抽出)

調査区は、平均して約50世帯を含む地区である。各標本調査区について、世帯名簿を作成し、この世帯名簿を基に標本調査区ごとに定められた抽出率を用いて系統抽出法により12世帯を抽出している。

(3) 「1日の生活時間」の調査日の選定

社会調の調査期日は10月20日としているが、「1日の生活時間」に関しては、平日および曜日ごとの結果を集計するため、標本調査区を無作為に8つのグループに分け、グループごとに9日間(10月13日から10月21日まで)のうち連続する2日間を調査日として選定している。

2.2 調査結果の推定方法

調査結果のうち、行動者数及び10歳以上人口は、国勢調査を基に推計した地域(都道府県)、男女、年齢階級別人口を基準人口とする比推定により算出している。また、延べ時間も同様に算出し、平均時間は、この延べ時間を対応する人口(または行動者数)の推定値で除して求めている。行動者率は、比推定値の百分比として算出している。平均時間を求める算式は次のようになる。

$$\begin{aligned} X \text{曜日における活動} Y \text{の平均時間} = \\ \frac{X \text{曜日における活動} Y \text{の延べ時間の比推定値}}{X \text{曜日における活動} Y \text{の人口の比推定値}} \end{aligned}$$

2.3 調査結果の標準誤差の推定方法

調査結果の推定値の標準誤差は、副標本によって推定している。副標本による標準誤差の推定は、第1段の標本抽出において独立で均等な大きさの数組の標本を抽出しておくのが本来であるが、社会調では、その近似として事後的に4組の副標本を設定している。事後的に4組の副標本を設定する方法は、都道府県、曜日(7曜日)ごとに調査区を配列し、1～4の番号を乱数で発生させて起番号とし

て、起番号から1～4の番号を順番に調査区に付与する。例えば、乱数で起番号「2」を選び、順番に2→3→4→1と組符号を調査区に付与する。この4組の副標本ごとに算出された推定値を用いて、次式により標準誤差を推定している。

$$\sigma = \sqrt{\frac{1}{4(4-1)} \sum_{i=1}^4 (X_i - X)^2}$$

σ : 標準誤差 i : 副標本の組番号

X_i : 第 i 副標本による平均時間などの推定値

X : 全標本による平均時間などの推定値

社会調は、生活時間の配分を調べるために、調査世帯を曜日別に振り分けて、調査しているので、曜日ごとに標準誤差を推定している。標準誤差の推定値としては、報告書に平日(月曜日～金曜日)、土曜日、日曜日の3区分と週全体を掲載している。

ここで、事後に副標本を設定することについて、事前に標本設計段階で副標本を設定する場合との比較で考える。社会調では、都道府県ごとに、調査区を一定の基準に基づいて配列し、系統抽出している。標本抽出の段階で同質の4組の副標本にする場合は、副標本にしない場合と比べると、4つの抽出起番号を無作為に選び、抽出間隔を4倍にして抽出することになる。これは、事後的に順序に従って組符号を振り分けて、副標本を設定することとほぼ同じことを行っている⁵⁾。ただし、抽出起番号を無作為抽出で選定するのか、系統抽出で選定するのかの違いはある。仮に、各副標本が同質でない場合は、標準誤差を過大推定することになる。しかし、各副標本における差異が同じように生じるならば、必ずしもばらつく要素にはならないと考えられる。なお、事前に副標本を設定している場合には、実際の調査で有効回答が得られなかった調査区が特定の副標本に偏る可能性が考えられるが、事後に設定する場合は、そのようなリス

クが回避される。

3. 検証方法

副標本による標準誤差のばらつきをシミュレーションによって検証する方法と、副標本によって推定する標準誤差をブートストラップ法による標準誤差の推定値によって検証する方法について述べる。

3.1 シミュレーションによる検証方法

副標本による標準誤差のばらつきについては、社会調のマイクロデータ(調査票情報)を用い、シミュレーションによって、副標本による標準誤差を幾通りにも推定して検証する。シミュレーションによって検証するのは、社会調が、労働力調査のように毎月の標準誤差を算出できず、1回しか算出することができないためである。

本稿では、曜日、都道府県ごとに4組の副標本を設定し、都道府県の4組の副標本を組み合わせ、曜日ごとに幾通りもの4組の副標本を作成することによってシミュレーションが可能になっている⁶⁾。

標準誤差の推定値としては、報告書の標準誤差の推定値と同様に平日、土曜日、日曜日の3区分と週全体とする。推定に使用しているのは、生活時間のうちの「趣味・娯楽」の時間である。「趣味・娯楽」とした理由は、「睡眠」や「食事」のように、たいていの人がある行動をとるような種類でなく、人によって選択が多様であり、時間のばらつきがちな種類の行動として選択した。

社会調と同様に、曜日ごとに事後的に分けた副標本ごとに生活時間の推定値を求める。副標本に分ける方法は、次のとおりである。

- (1) 調査区を都道府県、調査区の番号順に配列し、社会調と同様の方法(2.3節の方法)で4組の副標本に分ける。
- (2) 社会調における「調査区の抽出」と同じ配列とするために、調査区の並びに

大都市圏，都市階級を追加し，社会調と同様の方法で4組の副標本に分ける。

- (3) 調査区番号を使用せずに，都道府県ごとに調査区を無作為に並べ，組符号を順番に付与する方法で4組に分ける。
- (4) 系統的に番号を付与するのではなく，無作為に乱数で1～4の番号を付与する方法で4組に分ける。
- (5) 4組ではなく，副標本を6組，8組に分ける。副標本に分ける方法は(1)と同じとする。

(1)，(2)，(5)では，基本的には社会調で行っている組別に分ける方法をとっている。組符号を順番に付与して副標本に分けるのは，調査区の並びに一定の規則性があるため，それぞれの組に似通った標本調査区が振り分けられるためである。(3)，(4)においては，社会調で行っている方法ではなく，データを並び替える方法によって組別に分けている。(5)について，副標本の組数は多い方が標準誤差を安定的に推定できるものの，各副標本を同質にすることを考えるならば，副標本の組数は4組から10組までが望ましいとされている(浅井，1987)⁷⁾。しかし，4組という副標本の組数は少ないと考え，組数を増やした場合について行った。

以上の方法により，曜日ごと，都道府県ごとに副標本を設定し，都道府県の副標本を組み合わせ，曜日ごとに副標本を作成する。これらの副標本を用いて100回⁸⁾のシミュレーションを行い，曜日別，組別の生活時間の推定値を求め，曜日ごとの標準誤差を推定する。生活時間および標準誤差の推定には，それぞれ2.2節，2.3節の式を用いている。そして，平日，土曜日，日曜日の3区分と週全体に集約し，標準誤差の平均，標準誤差の標準偏差などを求める。この結果から，標準誤差がばらつくかどうか，ばらつきの大きさはどの程度かを評価し，ばらつきを大きくする要素についても考察する。

3.2 ブートストラップ法による検証方法

ブートストラップ法によって標準誤差の推定値を求め，標準誤差の推定値を比較することにより，副標本による標準誤差の推定値の大きさについて検証する。ブートストラップ法は，標本から再抽出することによってブートストラップ標本を作成して推定量を計算し，再抽出を何回も繰り返して，ブートストラップ標本間の推定量のばらつきから分散を推定する方法である(Efron, 1979)。このブートストラップ法によって標準誤差を推定するのは，標準誤差の算出式を導出することなく，複雑な理論式を大量の反復計算によって置き換えることができる手法であり，かつ安定的な標準誤差を算出できるとされているからである。

標本調査区を母集団とみなして，社会調の抽出方法に基づいて，復元を許した無作為抽出を1000回反復して推定値を求め，その推定値に基づいて，推定値 X の標準誤差を算出する。

$$\sigma = \sqrt{\frac{1}{(1000-1)} \sum_{i=1}^{1000} \{X(i) - X(\cdot)\}^2}$$

$$\text{ただし， } X(\cdot) = \frac{1}{1000} \sum_{i=1}^{1000} X(i)$$

σ ：標準誤差

$X(i)$ ：各回における平均時間の推定値

$X(\cdot)$ ：各回の平均時間の推定値の平均

ブートストラップ法においては，抽出回数を多くすると，標準誤差は安定する。ただし，特に何回以上必要と決まっているわけではない。そこで，どの程度の抽出回数で分散が安定するのかを試みたところ，抽出回数が多くなればなるほど安定性を増し，1000回ではほぼ安定的とみなせた。

ブートストラップ法による分散の推定では，無限母集団から無作為抽出した標本を想定しているため，有限母集団からの非復元無作為抽出した標本ではバイアスが生じるとされ，

バイアスを除去する方法が提案されている。その方法の1つである Rescaling Bootstrap 法 (Rao and Wu, 1988) を試したが、単純なブートストラップ法の値との差がない結果となった。そのため、バイアス調整を厳密にしなくとも、単純なブートストラップ法を適用して求めた標準誤差は検証に使えると判断した。

4. 生活時間における推定結果

シミュレーションによる検証方法で推定した結果に基づいて、標準誤差が副標本の構成によって、どのように変化しているのかを考察する。副標本による標準誤差の推定値については、ブートストラップ法によって推定した標準誤差との比較から考察する。

4.1 推定結果

「趣味・娯楽」における推定結果としては、100回のシミュレーションの結果を平均した平均時間と標準誤差を表に掲載している。標準誤差については、100回のシミュレーションにおける標準偏差、中位数、四分位範囲、最大と最小の倍率を示している。

(1) 4組による標準誤差

「趣味・娯楽」における週全体の平均時間 42.212 分、標準誤差 0.277 分、標準誤差率 0.66% である。標準誤差の標準偏差は、0.048 分となっている (表 1)。標準誤差は、週全体で小さく、日曜日で大きい結果となっている。標準誤差の標準偏差は、週全体、平日で小さく、土曜日、日曜日で大きくなっている。

(2) 大都市圏・都市階級で並び替えた 4組による標準誤差

「趣味・娯楽」における週全体の平均時間 42.220 分、標準誤差 0.324 分、標準誤差率 0.77% である。標準誤差の標準偏差は、0.055 分となっている (表 2)。この方法が社会調の方法に従っているので、本来の標準誤差に近い。(1)の結果と比べると、標準誤差および標準誤差の標準偏差ともに大きくなっている。大都

市圏・都市階級で並び替えることによって、標準誤差のばらつきは小さくなっていない。なお、大都市圏、都市階級の組み合わせではなく、都市階級のみで並び替えたとしても同様の結果であった。

(3) 調査区の並びを無作為にした 4組による標準誤差

「趣味・娯楽」における週全体の平均時間 42.212 分、標準誤差 0.287 分、標準誤差率 0.68% である。標準誤差の標準偏差は、0.044 分となっている (表 3)。(1)の結果と比べると、

表 1 生活時間の副標本による標準誤差 (4組)

		平日	土曜日	日曜日	週全体
平均時間 (分)		34.104	56.572	68.394	42.212
標準誤差	平均 (分)	0.360	0.371	0.551	0.277
	標準偏差 (分)	0.073	0.143	0.215	0.048
	中位数 (分)	0.358	0.363	0.548	0.274
	四分位範囲 (分)	0.104	0.233	0.277	0.067
	最大/最小 (倍)	3.414	9.222	8.096	2.298

注 1 : 平均時間、標準誤差ともに 100 回のシミュレーション結果の平均 (表 5 まで同じ)。

2 : 標準誤差の平均は全標本によって算出した値 (表 5 まで同じ)。

表 2 生活時間の副標本による標準誤差 (4組, 並び替え)

		平日	土曜日	日曜日	週全体
平均時間 (分)		34.117	56.567	68.388	42.220
標準誤差	平均 (分)	0.422	0.394	0.658	0.324
	標準偏差 (分)	0.081	0.175	0.247	0.055
	中位数 (分)	0.422	0.376	0.625	0.325
	四分位範囲 (分)	0.119	0.267	0.252	0.087
	最大/最小 (倍)	2.859	15.685	10.913	2.608

表 3 生活時間の副標本による標準誤差 (4組, 並びを無作為)

		平日	土曜日	日曜日	週全体
平均時間 (分)		34.103	56.573	68.398	42.212
標準誤差	平均 (分)	0.365	0.500	0.591	0.287
	標準偏差 (分)	0.067	0.199	0.233	0.044
	中位数 (分)	0.360	0.482	0.602	0.286
	四分位範囲 (分)	0.097	0.294	0.367	0.061
	最大/最小 (倍)	2.922	8.431	17.908	2.331

標準誤差は少し大きくなっている。標準誤差の標準偏差は、曜日によって異なるものの、ほぼ同じ程度の結果となっている。

(4) 無作為に4組に分けた場合の標準誤差

標本を乱数により4組に分けた場合については、各組の副標本の大きさが不均等⁹⁾になり、各副標本の同質性に問題があると考えられる。そのため、副標本によって標準誤差を求める場合に、この分け方を採用するのは適切でないと判断した。

(5) 副標本の組数を増やした場合の標準誤差

6組の副標本の場合、「趣味・娯楽」における週全体の平均時間42.194分、標準誤差0.273分、標準誤差率0.65%である。標準誤差の標準偏差は、0.036分となっている(表4)。(1)の結果と比べると、標準誤差は週全体、平日で小さく、土曜日、日曜日で大きくなっている。標準誤差の標準偏差は、総じて小さくなっている。8組の副標本の場合、標準誤差は日曜日を除いて大きくなっている。標準誤差の標準偏差は、小さくなっている(表5)。

(6) ブートストラップ法による標準誤差

ブートストラップ法による結果は、「趣味・娯楽」における週全体の平均時間42.499分、標準誤差0.297分、標準誤差率0.70%である(表6)。週全体の平均時間の標準誤差は、(1)との差はあまりない。

4.2 推定結果についての考察

副標本による標準誤差の推定において、標準誤差が大きくばらつくことは確認できた。ばらつきを小さくするために、副標本の同質性を高める方法として、データの並び替えを試みたものの、はっきりした結果は得られなかった。これは、都道府県別に分けているために市(区)町村番号順の並びと大都市圏・都市階級による並びにその差異がなかったためと推測される。

また、並びを無作為にする方法で4組に分けたところ、並びを規則的にした場合と結果

表4 生活時間の副標本による標準誤差(6組)

		平日	土曜日	日曜日	週全体
平均時間(分)		34.079	56.566	68.397	42.194
標準誤差	平均(分)	0.349	0.404	0.603	0.273
	標準偏差(分)	0.054	0.128	0.191	0.036
	中位数(分)	0.356	0.386	0.599	0.275
	四分位範囲(分)	0.086	0.174	0.280	0.054
	最大/最小(倍)	2.148	7.767	8.319	1.930

表5 生活時間の副標本による標準誤差(8組)

		平日	土曜日	日曜日	週全体
平均時間(分)		34.082	56.577	68.389	42.197
標準誤差	平均(分)	0.381	0.512	0.544	0.293
	標準偏差(分)	0.048	0.141	0.136	0.033
	中位数(分)	0.384	0.492	0.532	0.291
	四分位範囲(分)	0.072	0.160	0.189	0.045
	最大/最小(倍)	2.041	4.110	3.898	1.837

表6 生活時間のブートストラップ法による標準誤差

	平日	土曜日	日曜日	週全体
平均	34.211	57.517	68.925	42.499
標準誤差	0.382	0.547	0.625	0.297

にあまり差異がなかった。標本設計において調査項目に即した層別がなされていないこと、生活時間という調査項目の性格から、差異を生じさせなかったと考えられる。そのことが並びによって各副標本を同質化する効果がないようにみせているとも考えられる。

社会調において、副標本の組数は4組と少なく、ばらつきを生じさせる要素であるので、副標本の組数を増やして、標準誤差のばらつきを小さくする可能性を試した。組数を増やすことによってばらつきは小さくなるものの、土曜日では4組と8組がほぼ同じになるように、単純に組数を増やせば、ばらつきが小さくなるわけでもない。組数を増やしたとしても、同質性が保たれず、副標本の大きさ¹⁰⁾が十分確保されずに、標準誤差のばらつきが

大きくなる可能性もありえる。

ブートストラップ法による標準誤差の推定では、標準誤差のばらつきをみているわけではなく、副標本による標準誤差の推定値の妥当性を検証しようとするものである。副標本による結果の平均した値は、ブートストラップ法による結果にほぼ近似していると思われるので、推定値として妥当とみられる。

副標本による1回限りの推定では、最大と最小の倍率をみても、偶然に高い値、低い値になる可能性があるため、安定した標準誤差を推定するためには、複数回の推定値の平均か、または平均とほぼ近似している中位数によって妥当な結果が得られると考える。

5. 生活行動における検証

社会調では、生活時間のほかに生活行動についても調査している。生活行動については、調査期日(10月20日)までの1年間に活動したかどうかを調べている。生活行動についても検証し、その結果について述べることにする。

5.1 シミュレーションによる検証方法

社会調は、大規模標本調査ではあるが、生活時間の調査では、調査の性格上、曜日ごとに分ける必要があるため、標本の大きさの効果があまりないことが考えられる。生活行動の調査では、曜日に分ける必要がないので、生活行動の結果から標本の大きさの効果を確認する。シミュレーションの方法は、生活時間と同様とし、生活行動の「趣味・娯楽」の「スポーツ観覧(テレビ等は除く)」の行動者について算出する。生活時間での「趣味・娯楽」に合わせ、生活行動の中から比較的行動者が確保できる種目を選択した。生活時間の(1)、(5)の方法によって副標本に分け、(5)の方法における組数は、6組、8組、10組の3通りとする。

表7 生活行動の副標本による標準誤差

		4組	6組	8組	10組
行動者率(%)		19.423	19.423	19.417	19.412
標準誤差	平均(%)	0.162	0.160	0.181	0.173
	標準偏差(%)	0.066	0.048	0.048	0.044
	中位数(%)	0.153	0.162	0.179	0.171
	四分位範囲(%)	0.091	0.059	0.074	0.058
	最大/最小(倍)	12.145	6.096	4.181	3.265

注1：行動者率、標準誤差ともに100回のシミュレーション結果の平均。

注2：標準誤差の平均は全標本によって算出した値。

5.2 シミュレーションによる推定結果

4組の副標本による標準誤差の推定において、シミュレーション100回の平均した「スポーツ観覧(テレビ等は除く)」の行動者率は19.423%、標準誤差0.162%、標準誤差率0.83%である。標準誤差の標準偏差は、0.066%となっている。

4組、6組、8組、10組の副標本による標準誤差をみると、行動者率の標準誤差は、4組、6組の場合はほぼ同じ、8組、10組は大きくなっている。標準誤差の標準偏差は、6組、8組は同じ値であるものの、組数が増えることによって小さくなっている(表7)。

生活行動では、生活時間よりも各副標本の大きさは確保されるので、標準誤差の数値にも影響せず、ばらつきは小さくなっていると考えられる。

6. おわりに

副標本による標準誤差の計測は、1つの有効な方法であるものの、生活時間の推定結果からみると、標準誤差は安定的ではない。副標本の組数を増やした場合の標準誤差をみると、組数を多くするだけではなく、副標本の同質性を確保するためにも、ある程度の標本の大きさを考慮して、副標本にする必要性があることが確認できる。

数年に1回の調査においては、副標本による標準誤差が妥当な結果であることを本稿のシミュレーションのような実験によってしか

確認することはできない。そのため、実務的には、過去の調査結果と比較する、データを並び替えて試算した結果¹¹⁾と比較する、ブートストラップ法のようなサンプリング手法による試算結果と比較することによって、妥当な結果であるかを確認するのが望ましい。また、ブートストラップ法によって標本誤差を評価するのも一つの方法と考えられる。

最後に、本稿では、社会生活基本調査の調

査結果から副標本による標本誤差の計測について検証したが、この結果は社会生活基本調査の調査項目の性格に大きく依存していることも考えられる。副標本の組数、副標本の大きさ、調査項目によって、それぞれの調査での表れ方は異なると思われるので、副標本によって標本誤差を計測している他の大規模標本調査においても検証をすることは重要と考える。

注

- 1) 副標本による標本誤差の推定については、Wolter(1985)などを参照のこと。
- 2) 社会生活基本調査、就業構造基本調査、労働力調査の副標本の組数は、それぞれ4組、6組、8組である。
- 3) 労働力調査の副標本による標本誤差のばらつきは、副標本の大きさが十分に大きくないこと、副標本の組数が少ないこと、実際に調査された結果としての各副標本の値には非標本誤差が含まれていることから生じるばらつきであると考えられる。また、副標本による標本誤差は、ブートストラップ法による標本誤差や理論式に基づいた標本誤差に比べると、少し大きく計測されている。
- 4) 社会生活基本調査において調査する事項の性格上、標本調査区は、社会施設・大きな病院のある調査区など一部を除いた調査区の中から抽出している(総務省統計局, 2003)。
- 5) 事後に作成する副標本は、独立に抽出された標本ではない。そのため、副標本による分散は、不偏推定量ではなく、過大推定となっている。
- 6) 全国の分散を推定するのに、都道府県ごとに4組の副標本によって分散を求め、その分散を加重平均して推定する方法はとっていない。この方法では、都道府県ごとの分散は一通りしか得られず、シミュレーションによる実験を行うことができないためである。
- 7) 副標本の組数について、土屋(2009)は、安定した分散を推定するためには10組以上が必要としている。
- 8) 労働力調査での副標本による標準誤差の推定では、12か月分の標準誤差から標準誤差のばらつきを確認している。そして、12か月分の標準誤差の平均が理論式による標準誤差やブートストラップによる標準誤差よりも大きく計測されているが、平均することによって安定的な結果が得られている。そのことからシミュレーションの回数は、12回以上あれば有効と考えたが、多くして100回とした。
- 9) 無作為による4組に分ける方法では、各副標本の大きさは均等にならない。日曜日の副標本の大きさを例にとると、調査区数で976, 937, 967, 927(全調査区 3807)のように差異が生じる。他の分ける方法でも差異は生じる場合があるものの、ほぼ均等とみなせる。
- 10) 社会調の標本の大きさは、生活時間については、平日約138,400人、土曜日約115,000人、日曜日約114,500人、週全体約367,900人である。生活行動については、約185,900人である。1副標本当たりの大きさは、4組の場合それぞれ4分の1で、平日約34,600人、土曜日約28,700人、日曜日約28,600人、週全体約92,000人、生活行動約46,500人である。
- 11) データを並び替える方法では、その方法によって分散を過大推定することに留意する必要がある。

参考文献

- [1] 浅井晃(1987)『調査の技術』日本科学技術連盟出版社。
- [2] 栗原由紀子(2010)「社会生活基本調査マイクロデータにおける平日平均統計量と標本誤差の計

- 測』『統計学』第99号, 20～35頁.
- [3] 斎藤金一郎・浅井晃・大沢豊訳(1964)『調査における標本設計』日本科学技術連盟出版社 (Deming, W.E. (1960) *Sample Design in Business Research* John Wiley & Sons, Inc. の全訳).
 - [4] 総務省統計局 (2003) 「平成13年社会生活基本調査報告」.
 - [5] 総務省統計局 (2005) 「労働力調査 標本設計の解説」.
 - [6] 高橋雅夫・臼井彩子 (2005) 「平成13年社会生活基本調査における標本の代表性と調査結果の推定について」『統計研究彙報』第62号, 23～70頁.
 - [7] 竹内啓編 (1989) 『統計学辞典』東洋経済新報社.
 - [8] 土屋隆裕 (2009) 『概説 標本調査法』朝倉書店.
 - [9] 馬場康維・土屋隆裕・中村好宏・山崎伸彦 (1996) 「ブートストラップによる標準誤差の推定の試み」『第10回計算機統計学会大会論文集』, 68～71頁.
 - [10] 馬場康維・土屋隆裕・中村好宏 (1997) 「個票データの利用による推定精度の評価実験」『第65回日本統計学会講演報告集』, 156～157頁.
 - [11] 馬場康維・土屋隆裕・中村好宏・小林良行 (1997) 「労働力調査におけるブートストラップ法の利用」『第65回日本統計学会講演報告集』, 224～227頁.
 - [12] 馬場康維・土屋隆裕 (1998) 「ブートストラップ推定 — 理論と実用性 — 」『第66回日本統計学会講演報告集』, 42～43頁.
 - [13] 標本誤差推計研究会編 (1998) 『標本誤差の推計方法 — 最新時代の理論と実証 — 』統計情報研究開発センター.
 - [14] 古橋正宏・岩永琢磨 (1991) 「労働力調査の標本誤差」『統計局研究彙報』第49号, 37～49頁.
 - [15] 松田芳郎・伴金美・美添泰人 (2000) 『講座ミクロ統計分析 ミクロ統計の集計解析と技法』日本評論社.
 - [16] 村田磨理子 (1998) 「人工データによる誤差推定方法の比較」『ESTRELA』No. 48, 20～27頁.
 - [17] Cochran, W.G. (1977) *Sampling Techniques*, Third Edition, John Wiley, New York.
 - [18] Efron, B. (1979) “Bootstrap Methods: Another Look at the Jackknife”, *The Annals of Statistics*, Vol. 7, No. 1, pp.1-26.
 - [19] Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*, Chapman & Hall, London.
 - [20] Rao, J.N.K. and Wu, C.F.J. (1988) “Resampling inference with complex survey data”, *Journal of the American Statistical Association*, 83, pp.231-241.
 - [21] Wolter, K.M. (1985) *Introduction to Variance Estimation*, Springer-Verlag, New York.

Estimation of Sampling Errors by using Sub-Samples

Kozo YAMAGUCHI*

Summary

Some official statistical surveys estimate sampling errors by using sub-samples. This method has advantages in that sampling errors can be easily estimated, even in surveys with a complex sample design. However, the method has disadvantages in that the sampling errors are varied.

The Survey on Time Use and Leisure Activities is a statistical survey carried out at several year intervals with a large sampling unit. In this paper, we study whether the sampling errors by using sub-samples in this survey are varied, and the degree of the variation, by using a simulation technique. As a result, we verify that the sampling errors by using sub-sample varied. In addition, we present a method of estimating a stable sampling error.

Key Words

Survey on Time Use and Leisure Activities, Sampling Error, Sub-Sample, Bootstrap

* Statistical Research and Training Institute, Ministry of Internal Affairs and Communications
2-11-16 Izumi-cho, Kokubunjishi, Tokyo 185-0024 Japan
e-mail : k3.yamaguchi@soumu.go.jp

編集委員会からのお知らせ
機関誌『統計学』の編集・発行について

編集委員会

本年9月より、新しい規定にもとづいて、「研究論文」と「報告論文」が設定されました。皆様からの積極的な投稿をお待ちしております。

1. 投稿は、常時、受け付けています。なお、書評、資料および海外統計事情等については、下記の[注記2]をご確認下さい。
2. 次号以降の発行予定日は、
第112号：2017年3月31日、第113号：2017年9月30日です。
3. 投稿に際しては、新規定にもとづく「投稿規程」、「執筆要綱」、および「査読要領」などをご熟読願います。最新版は、学会の公式ウェブサイトをご参照下さい。
4. 原稿は編集委員長（下記メールアドレス）宛にお送り下さい。
5. 原稿はPDF形式のファイルとして提出して下さい。また、紙媒体での提出も旧規程に準拠して受け付けます。紙媒体の送付先は編集委員長宛にお願いします（住所は会員名簿をご参照下さい）。
6. 原則として、すべての投稿原稿が査読の対象となります。
7. 通常、査読から発刊までに要する期間は、査読が順調に進んだ場合でも、2ヶ月間程を要します。投稿にあたっては十分に留意して下さい。

編集委員会、投稿応募についての問い合わせは、
下記メールアドレス宛に連絡下さい。
また、編集委員長へのメールアドレスも下記になります。

editorial@jsest.jp

編集委員長 朝倉啓一郎（流通経済大学）
副委員長 藤井輝明（大阪市立大学）
編集委員 橋本貴彦（立命館大学）
前田修也（東北学院大学）
山田 満（東北・関東支部所属）

[注記1] 『統計学』の定期刊行に努めておりますので、できるかぎり早期のご投稿をお願いします。112号（2017年3月31日発行予定）への掲載を想定した場合、「研究論文」と「報告論文」の原稿は、2017年1月初旬を目途として、それまでにご投稿ください。

[注記2] 書評、資料および海外統計事情等について、執筆、推薦、および依頼等をお考えの会員がおられましたら、企画や思いつきの段階で結構ですので、できるだけ早い段階で、編集委員会にご一報下さい。

以上

編集後記

研究成果を投稿下さいました執筆者の皆様、査読に関わって下さいました皆様、そして、書評の依頼をお引き受け下さいました皆様に、心からお礼申し上げます。とくに、本号は、本年9月からスタートした新規定にもとづく編集作業でもありましたので、関係する多くの皆様のご支援を頂くことで、発行することが出来ました。編集委員一同、重ねて感謝申し上げます。

さて、次号112号からは、通常の論文に加えて、「『統計学』創刊60周年記念特集論文」の掲載が開始される予定です。楽しみにお待ちしております。

編集委員会では、機関誌『統計学』を充実させていくために、皆様からの率直なご意見と、そして、研究成果の積極的なご投稿をお待ちしております。今後ともよろしくごお願い申し上げます。

（朝倉啓一郎 記）

執筆 者 紹 介 (掲載順)

高橋雅夫 (総務省統計局) 高部 勲 (総務省統計局)
山口幸三 (総務省統計研修所) 宮川幸三 (立正大学経済学部)
居城 琢 (横浜国立大学国際社会科学研究院) 櫻本 健 (立教大学経済学部)
大西 広 (慶應義塾大学経済学部)

支 部 名

事 務 局

北 海 道	062-8605 札幌市豊平区旭町 4-1-40 北海学園大学経済学部 (011-841-1161)	水野谷 武志
東 北・関 東	980-8511 仙台市青葉区土樋 1-3-1 東北学院大学経済学部 (022-721-3417)	前 田 修 也
関 西	567-8570 茨木市岩倉町 2-150 立命館大学経営学部 (072-665-2090)	田 中 力
九 州	870-1192 大分市大字旦野原 700 大分大学経済学部 (097-554-7706)	西 村 善 博

編 集 委 員

朝倉啓一郎 (東北・関東) [長] 藤井輝明 (関 西) [副]
前田修也 (東北・関東) 橋本貴彦 (関 西)
山田 満 (東北・関東)

統 計 学 No.111

2016年9月30日 発行	発行所	経 済 統 計 学 会 〒112-0013 東京都文京区音羽1-6-9 音羽リスマチック株式会社 TEL/FAX 03 (3945) 3227 E-mail: office@jsest.jp http://www.jsest.jp/
	発行人	代表者 西 村 善 博
	発売所	音羽リスマチック株式会社 〒112-0013 東京都文京区音羽1-6-9 TEL/FAX 03 (3945) 3227 E-mail: otorisu@jupiter.ocn.ne.jp 代表者 遠 藤 誠

STATISTICS

No. 111

2016 September

Articles

- An estimation of establishment birth and death rates based on the Economic Census
..... Masao TAKAHASHI, Isao TAKABE (1)

Short Articles

- Estimation of Sampling Errors by using Sub-Samples
..... Kozo YAMAGUCHI (17)

Book Reviews

- Ichiro ASARI and Eiji DOI, *The Theory and Practices of Inter-Regional Input-Output Analysis*, Nippon Hyoron sha, 2016
..... Kozo MIYAGAWA (27)
- Takayuki YAMASHITA ed., *Handbook of Regional Economic Analysis: Regional revitalization learned from Shizuoka Model*, Koyo Shobo, 2016
..... Taku ISHIRO (32)
- Jie LI, *Introductory GDP statistics and input-output analysis*, University Education Press, 2016
..... Takeshi SAKURAMOTO (38)
- Tadasu MATSUO and Takahiko HASHIMOTO, *An Introduction to Tomorrow's Marxian Economics*, Chikumashobo, 2016
..... Hiroshi ONISHI (43)

Activities of the Society

- The 60th Session of the Society of Economic Statistics (46)
- Regulation of the Editorial Committee, Prospects for the Contribution to the Statistics (72)

JAPAN SOCIETY OF ECONOMIC STATISTICS
