

全国物価統計調査における分位点推定値の標本誤差の検討

樋田 勉*

計量経済学研究室

A study on the sampling errors of quantile estimates in National Survey of Prices

Tsutomu TOIDA

Econometrics

概 要

本稿では、Funaoka, Saigo, Sitter, and Toida (2006) が考案した層別多段抽出のためのベルヌーイ・ブートストラップ法 (BBE) を利用し、平成9年、14年全国物価統計調査における、分位点推定値の標準誤差の推定を行った。本稿で行った計算から、分位点推定値の標準誤差は非常に小さく推定精度は高いこと、および、平成9年、14年の間で推定精度に大きな変化はないことが示された。

Abstract

In this article, we estimate the standard errors of quantile estimates of price distributions in the 1997 and 2002 National Survey of Prices applying the Bernoulli bootstrap method for stratified multistage sampling proposed by Funaoka, Saigo, Sitter, and Toida (2006). Our

*群馬大学社会情報学部 〒371-8510 群馬県前橋市荒牧四丁目2番地 toida@si.gunma-u.ac.jp

本稿は、舟岡史雄 信州大学教授、西郷浩 早稲田大学教授と筆者が行っている平成9年、14年全国物価統計調査、商業統計調査の個票データを利用する共同研究において、筆者が分析を担当した部分である。両調査の個票利用については、指定統計調査の目的外使用申請を行い、その利用が官報(平成18年3月28日第4305号、総務省告示第176号、総務省告示第177号)で告示されている。

本稿をまとめるにあたり、舟岡教授、西郷教授から大変貴重なご意見・ご指摘をいただきました。また、目的外使用申請にあたり総務省統計局物価統計室のご助力をいただきました。心からお礼申し上げます。筆者は文部科学省科学研究費補助金(課題番号16203014, 18530156)、群馬大学から研究助成を受けています。記して謝意を表します。

findings indicate that the standard errors of quantile estimates are very small, so the estimates are thought to be highly accurate. And we also find that there are no remarkable difference in standard errors between 1997 and 2002.

1 はじめに

全国物価統計調査は、総務省が5年ごとに実施する指定統計調査(108号)である。第1回調査は、昭和42年に実施され、基本的に5年ごとに調査が実施されてきた。調査の目的は、「国民の消費生活において重要な支出の対象となる商品の価格及びサービスの料金並びにこれらを取り扱う店舗の経営形態や立地環境など価格決定に関する様々な要素を幅広く調査し、価格の店舗間格差、銘柄間格差、地域間格差など価格差の実態を解明し、物価に関する基礎資料を得ること」である。

全国物価統計調査の調査方法は、平成9年に大幅に変更された。平成9年以降の調査では、層別多段抽出によって調査店舗の選定が行われる。抽出フレームには、経済産業省が実施する商業統計調査の商業準備調査名簿が利用される。

全国物価統計調査の調査事項は、店舗の名称、業態、競合店の有無等、小売店舗の基本的属性と、300銘柄以上の財・サービスの小売価格や料金である。従って、全国物価統計調査の結果から、物価の横断面的な分析が可能である。

平成9年全国物価統計調査を利用する価格分布の研究についての簡単なサーベイが樋田(2006)にある。個票データを利用した研究としては、菅(2004)、樋田(2006)、Funaoka, Saigo, Sitter, and Toida(2006)がある。菅(2004)は、消費者物価指数の作成方法に関するいくつかの批判に対して、平成4年、平成9年の個票データを利用し、客観的な検証を行っている。樋田(2006)は、全国物価統計調査と商業統計調査の個票データを利用し、厳密な銘柄管理が行われている指定商標銘柄(「日清カップヌードル 醤油味(日清食品, 76g)」のように商標・規格・容量などが定められている銘柄)の価格分布の分析を行った。そして、指定商標銘柄の価格分布の多くは、複数の峰を持つ多峰型分布であり、小売店舗の業態、販売戦略、店舗の立地環境等、小売店舗の経営戦略や、経営環境の相違が、価格分布が多峰となる要因であることを示している。一方、全国物価統計調査は標本調査であり、表章された平均価格や分位点には標本誤差が含まれる。Funaoka, et al.(2006)は、層別多段抽出の各段階で単純無作為抽出が利用されているとき、第1段の抽出率が高い場合でも、利用することができるブートストラップ法を考案した。そして、平成9年全国物価統計調査の個票データを利用して標本誤差の評価を行い、標本誤差が十分小さいことを示している。

平成14年調査は、大規模な調査方法の変更が行われてから2回目の調査となる。従って、平成9年、14年調査の標本誤差が安定的で、精度が高い結果が得られているか検証することは、統計作成者、利用者双方にとって有益である。

本稿の目的は、平成9年、14年全国物価統計調査の個票データを利用し、分位点推定値の精度が安

定的であるか検証することである。標本誤差の計算には、Funaoka, et al. (2006) のベルヌーイ・ブートストラップ法 (BBE) を利用する。また、全国物価統計調査への適用方法についても、Funaoka, et al. (2006) の方法を採用する。従って、本稿の貢献は、Funaoka, et al. (2006) の方法が、平成14年全国物価統計調査にも適用可能であることを示し、両年の精度を検討した部分に限られる。

以下の構成は次の通りである。第2節では、全国物価統計調査の標本設計と、BBEの適用方法をFunaoka, et al. (2006) に従って示す。そして、平成9年、14年両調査で調査が行われた銘柄について、価格分布の分位点推定値と標本誤差を比較検討する。第3節では分析結果をまとめるとともに、今後の展望について述べる。

2 推定精度の検討

2.1 全国物価統計調査の標本設計

全国物価統計調査の調査店舗選定は、売り場面積が450㎡以上の店舗を大規模店舗、450㎡未満の店舗を小規模店舗として、それぞれ層別多段抽出によって行われる。抽出フレームは、同年実施の経済産業省商業準備調査名簿である。大規模店舗の抽出では、1次抽出単位は市町村、最終抽出単位は店舗である。また、小規模店舗の抽出では、1次抽出単位は市町村、2次抽出単位は調査地区、最終抽出単位は店舗である。

1次抽出単位の抽出では、都道府県ごとに、市町村を経済圏などによりブロックに分け、ブロック内の市町村を人口階級別に層化する。層の数は全国で500以上となる。各層の市町村は、市町村に所在する小売店舗の年間商品販売額合計額の大きい方から配列され、人口階級によって定められた抽出率により、系統抽出される。表1に、平成9年、14年調査における市町村の抽出率を示した。表1から、市部の抽出率が高く、人口が4万人未満の町村の抽出率は低いことがわかる。

表1 標本設計とリサンプリング法(1)

人口階級	平成9年		平成14年		リサンプリング法 (平成9年、14年共通)	
	調査市町村数	抽出率	市町村数	調査市町村数	抽出率	市町村数
10万以上の市	221	1/1	221	228	1/1	229
5～10万未満の市	179	2/3	220	124	1/2	217
5万未満の市	80	1/3	224	111	1/2	226
4万以上の町村	4	1/5	32	20	2/3	34
4万未満の町村	187	1/15	184	153	1/15	2524

(注) 抽出された市町村数が少ないためにブロック毎、都道府県毎にBBEを行えない場合は、地理的に近接する都道府県、ブロックを統合して1つの層とする。

表2は、2次抽出単位、3次抽出単位の設定と抽出方法についてまとめたものである。

大規模店舗の抽出では、調査市町村に存在するすべての店舗が調査される。従って、実質的には層別1段抽出である。

小規模店舗の調査地区は、商業統計調査の調査区を用いて作成される。調査地区の作成方法と、地区に含まれる店舗数は、市町村の人口規模によって異なる。そして、市町村の人口規模によって定められた抽出地区数が、系統抽出される。抽出された調査地区に所在する店舗は、年間商品販売額の大きい順に配列され、40店舗が抽出される（詳細な調査方法については、総務庁統計局（2000）、総務省統計局（2005）を参照されたい）。

表2 標本設計とリサンプリング法(2)

	全国物価統計調査の調査設計		リサンプリングにおける処理
大規模店舗	調査市町村内に所在するすべての店舗を調査する。		悉皆抽出
小規模店舗	調査地区の抽出	人口10万以上の市	調査地区の設定 調査区の属性（商業集積地、住宅地及びその他）は無視し、各調査地区内の店舗数は90であるとする。市町村に、複数の調査地区が設定されているが、抽出調査地区数が1の場合、地理的に近接する市町村と統合する。 調査地区の抽出 単純無作為抽出と同等であると見なす。
		人口10万未満の市、町村	調査地区の設定 商業統計調査の調査区を組み合わせ、店舗数が70以上になるように調査地区を設定する。 調査地区の抽出 市町村の人口階級ごとに定められた調査地区数を系統抽出する。
	店舗の抽出	調査地区内に所在する店舗を年間商品販売額の大きい順に配列し、40店舗を系統抽出する。	単純無作為抽出と同等であると見なす。店舗の正確な抽出率は利用できないため、抽出率は1/2と仮定する。

2.2 BBE法の適用方法

全国物価統計調査の標本設計には系統抽出が利用されており、厳密な精度評価を行うことはできない。系統抽出が単純無作為抽出と同等であると仮定しても、1段目の抽出率が高い場合の層別多段抽出における分位点推定値の精度評価には、標本調査の領域で頻繁に利用されるような公式やリサンプリング法を利用することは難しい。

Funaoka, et al. (2006) は、各段階の抽出に単純無作為抽出が利用されている場合、1段目の抽出率が高い層別多段抽出にも容易に適用することができるブートストラップ法（Short cut BBE および General BBE）を開発し、short cut BBE を用いて、平成9年全国物価統計調査の分位点推定値の標準誤差の計算を行った。平成14年調査は、平成9年調査を踏襲しており、Funaoka, et al. (2006) の方法で、標準誤差を計算できる。以下、Funaoka, et al. (2006) に従い、リサンプリング法を適用する方法について述べる。

表1, 2に、リサンプリング法適用のために必要となる、各段階における単純化や仮定についてまとめてある。全国物価統計調査の各段階の抽出では系統抽出が行われるが、系統抽出が単純無作為抽出

出と同等であると見なす。また、全国物価統計調査では、全国に500以上の層が設定され、層ごとに市町村の抽出が行われる。層によっては、抽出された市町村数が少ないために BBE 法を適用できない場合がある。このような場合は、近接する都道府県、ブロックを統合して新しい層を作成する。この処理により、平成9年では約280、14年では約260の層に統合される。また、町村は1段目の抽出率が小さいため、BWR (Bootstrap With Replacement) 法によってリサンプリングを行う (BWR については、Shao and Tu (1995) を参照されたい)。

小規模店舗の調査地区の抽出では、リサンプリング法を適用するため、さらにいくつかの単純化を行う。1つ目は、市町村に複数の調査地区が設定されているにもかかわらず、抽出された調査地区が1の市町村の処理である。このような市町村があると、リサンプリング法を適用することができないため、近接する市町村と統合する。2つ目は、調査地区の属性は無視する。3つ目は、各調査地区における店舗の抽出率は1/2と単純化する。これらの仮定や単純化が標準誤差の推定に与える影響は無視できる程度であると見なし、BBE を適用する。BBE についての簡単な解説を補論に記す。

2.3 BBE 法による標準誤差の推定

表3は、平成9年、14年全国物価統計調査の個票データを利用して計算した、うるち米、即席めん、ヨーグルト、ビール、ドリンク剤、背広服、電気冷蔵庫、ビデオカメラの分位点推定値と BBE 法による標準誤差である。

うるち米では、大規模店舗の標準誤差は全体的に平成9年より平成14年が小さい。平成9年の25%点、50%点の標準誤差は、他の分位点の標準誤差と比べて大きい、分位点推定値の0.2%程度である。小規模店舗でも、標準誤差の分位点推定値に対する相対的な大きさは、平成9年の25%点における約0.2%が最大である。

即席めん、ヨーグルト、ビール、ドリンク剤では、半数以上の分位点推定の標準誤差が0である。このほかの銘柄でも、標準誤差が0の分位点が多い。このことは、指定商標銘柄の価格分布の多くが離散的であることによる (樋田 (2006) を参照されたい)。

背広服では、大規模店舗の標準誤差のほとんどが0である。小規模店舗の標準誤差は、大規模店舗に比べて大きい、分位点推定値に対する相対的な大きさは、最大でも3%未満であり、十分な推定精度が得られていることがわかる。

電気冷蔵庫、ビデオカメラでは、分位点推定値に対する標準誤差の相対的な大きさは、すべて1%未満である。従って、ここでも十分な推定精度が得られていることが確認できる。

以上の結果から、各銘柄に共通して、分位点推定値の標準誤差は非常に小さいことが示された。また、平成9年、14年の比較でも、推定精度に大きな変化は見られないことが示された。

表3 分位点推定値と標準誤差

品目・銘柄	店舗規模	実施年	分位点	10	25	50	75	90
うるち米 5kg「コシヒカリ」(魚沼産を除く)	大規模	9年	推定値	2258	2468	2655	2919	3035
			標準誤差	3.14	5.18	5.49	0.00	3.52
		14年	推定値	2079	2184	2404	2614	2814
			標準誤差	0.00	0.00	2.49	3.76	0.00
	小規模	9年	推定値	2394	2552	2783	2991	3150
			標準誤差	2.28	5.26	2.12	0.32	5.90
		14年	推定値	2184	2394	2604	2782	2940
			標準誤差	4.16	2.66	0.00	0.00	2.07
即席めん 「カップヌードル 醤油味」(日清食品)	大規模	9年	推定値	105	124	145	150	150
			標準誤差	0.00	0.00	0.07	0.00	0.00
		14年	推定値	101	103	123	144	150
			標準誤差	0.00	0.17	0.00	0.45	0.00
	小規模	9年	推定値	134	147	150	155	158
			標準誤差	0.00	0.70	0.00	0.94	0.00
		14年	推定値	123	144	150	150	152
			標準誤差	0.00	0.00	0.00	0.00	0.00
ヨーグルト 「明治 プルガリアヨーグルト LB81」(明治乳業), 500g	大規模	9年	推定値	166	187	208	250	260
			標準誤差	0.00	0.00	0.00	0.00	0.00
		14年	推定値	165	176	199	238	249
			標準誤差	0.00	0.00	3.39	0.67	0.00
	小規模	9年	推定値	176	208	242	250	260
			標準誤差	0.63	0.00	1.43	0.00	0.00
		14年	推定値	165	186	207	249	260
			標準誤差	0.32	0.00	0.00	0.00	0.00
ビール24缶 「アサヒ スーパードライ」, 24缶	大規模	9年	推定値	4179	4179	4494	4905	5250
			標準誤差	1.68	0.00	0.00	8.66	8.08
		14年	推定値	4074	4179	4179	4536	4964
			標準誤差	0.00	0.00	0.00	16.24	0.00
	小規模	9年	推定値	4662	5000	5355	5494	5496
			標準誤差	14.87	0.00	18.23	0.00	0.96
		14年	推定値	4200	4800	4989	5250	5493
			標準誤差	15.75	0.00	0.00	0.00	0.00
ドリンク剤 「リポビタンD」(大正製薬), 10本	大規模	9年	推定値	1020	1029	1134	1134	1239
			標準誤差	0.15	0.00	0.00	0.00	0.00
		14年	推定値	997	1005	1029	1029	1134
			標準誤差	0.69	0.00	0.00	0.00	0.00
	小規模	9年	推定値	1029	1050	1155	1260	1365
			標準誤差	1.14	4.62	0.47	0.00	0.00
		14年	推定値	1000	1029	1134	1470	1533
			標準誤差	0.00	0.00	0.00	2.80	0.00
背広服 秋冬物, シングル上下, 並型, 中級品	大規模	9年	推定値	20790	30450	37590	41790	61950
			標準誤差	0.00	0.00	0.00	0.00	0.00
		14年	推定値	10290	18900	30450	32216	51450
			標準誤差	0.00	272	0.00	0.00	0.00
	小規模	9年	推定値	34650	41790	57750	72450	89250
			標準誤差	939	0.00	582	199	1423
		14年	推定値	29400	39900	52500	72450	89250
			標準誤差	273	314	263	192	498
電気冷蔵庫 平成9年 ナショナル「NR-D41C1」 平成14年 ナショナル「NR-E40G1」	大規模	9年	推定値	157290	186900	218400	228900	239400
			標準誤差	0.00	672	0.00	0.00	0.00
		14年	推定値	104790	113400	122850	141540	155400
			標準誤差	0.00	0.00	0.00	644	500
	小規模	9年	推定値	194250	204540	225250	236565	249900
			標準誤差	553	1285	418	285	0.00
		14年	推定値	141750	155400	165900	176400	197400
			標準誤差	539	0.00	0.00	7.42	0.00
ビデオカメラ 平成9年 ソニー「DCR-TRV7」 平成14年 ソニー「DCR-TRV50」	大規模	9年	推定値	173040	176400	186900	197400	207900
			標準誤差	173	0.00	0.00	0.00	0.00
		14年	推定値	130767	138000	151935	165690	172200
			標準誤差	304	814	447	0.00	61.1
	小規模	9年	推定値	176400	186900	198000	209790	222075
			標準誤差	0.00	0.00	394	127	216
		14年	推定値	134400	144900	155400	167790	187215
			標準誤差	52.0	769	98.0	534	667

3 おわりに

本稿では、Funaoka, et al. (2006) の BBE 法を利用して、平成 9 年、14 年全国物価統計調査の価格分位点推定値の精度の比較を行った。本稿で取り上げたすべての銘柄で、分位点推定値の精度は非常に高いことが示された。また、平成 9 年、14 年の比較でも、標準誤差が大きく変化することはなく、精度は安定的であった。本稿で取り上げた銘柄は、全国物価統計調査の調査銘柄の一部であるが、他の銘柄でも同様の傾向である。分位点の標準誤差が非常に小さくなる原因の 1 つは、価格分布が実際には離散的な分布であり、一部の代表的な価格で非常に多くの店舗が販売していることである。

全国物価統計調査報告書では、店舗の業態、ディスカウント販売の有無など様々な店舗属性によって店舗が層化され、分位点推定値が示されている。本稿では取り上げていないが、層化を行う際、層に属する店舗数が極端に少なくならない限り、分位点推定値の精度は高い。また、平成 9 年、14 年の比較でも安定的な結果が得られており、表章された結果の誤差は小さいと考えてよい。

近年の物価のデフレ傾向の中で、研究者、政策担当者、金融市場関係者を始め、様々な方面から消費者物価指数や物価への関心は高まっており、特に消費者物価指数の精度についての議論は多い。全国物価統計調査は、5 年に一度の調査ではあるが、物価に関する非常に広範な情報を集めており、消費者物価指数に対して提起される様々な議論に答えることができる情報を含んでいる。本稿で、全国物価統計調査の分位点推定値の精度が非常に高いことを示した意義は大きいと考えられる。一方、本稿では、消費者物価指数の精度や、物価構造に関する検討は取り上げていない。これらの問題について、全国物価統計調査、小売物価統計調査、POS データ等の統計データに基づく検討を行うことが今後の課題である。

補 論

ここでは、本稿で利用した Funaoka, Saigo, Sitter, and Toida (2006) の層別 3 段抽出のための BBE 法 (Short cut BBE) について簡単に記しておく。

層別 3 段抽出

有限母集団は、互いに交わらない H 層に分割され、各層には N_1, \dots, N_H 個の 1 次抽出単位 (PSU, Primary sampling units) が含まれる。母集団全体では、 $N = \sum_{h=1}^H N_h$ の PSU がある。1 段目の抽出は、層ごとに独立な非復元単純無作為抽出 (SRSWOR, Simple random sampling without replacement) によって行われ、 n_1, \dots, n_H の PSU が抽出される。抽出される PSU の総数は $n = \sum_{h=1}^H n_h$ である。

第 h 層の i 番目の 1 次抽出単位 PSU_{hi} には、 M_{hi} 個の 2 次抽出単位 (SSU, Secondary sampling units) が含まれる。1 段目の抽出で PSU_{hi} が抽出されると、2 段目の抽出では、 M_{hi} 個の SSU から m_{hi}

個が、SRSWOR により抽出される。

PSU_{hi} の j 番目の 2 次抽出単位 SSU_{hij} には、 L_{hij} 個の最終抽出単位 (USU, Ultimate sampling units) が含まれる。2 段目の抽出で SSU_{hij} が抽出されると、3 段目の抽出では、 L_{hij} 個の USU から l_{hij} 個が、SRSWOR により抽出される。標本に含まれる USU のサイズは、 $\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} l_{hij}$ である。また、1 段目の抽出率は、 $f_{1h} = n_h / N_h$ 、2 段目の抽出率は $f_{2hi} = m_{hi} / M_{hi}$ 、3 段目の抽出率は $f_{3hij} = l_{hij} / L_{hij}$ である。

Short cut BBE

層別 3 段抽出のもとで、特性値の母集団分位点を推定し、分散の計算を行う。母集団分位点の推定は通常の方法で行う。分散を計算するためのリサンプリング標本の構成は、以下の 3 段階によって行う。

- (1) 第 h 層の i 番目の 1 次抽出単位 PSU_{hi} を、確率 $p_h = \sqrt{1 - \frac{1 - f_{1h}}{1 - n_h^{-1}}}$ で、ブートストラップ標本に含め、(2) に進む。 PSU_{hi} をブートストラップ標本に含めない場合は、第 h 層の n_h 個の PSU から、1 つの PSU をランダムに取り出しブートストラップ標本に含める。
- (2) (1) で PSU_{hi} をブートストラップ標本に含めた場合、 PSU_{hi} の j 番目の 2 次抽出単位 SSU_{hij} を確率 $q_{hi} = \sqrt{1 - \frac{f_{1h}}{p_h^{-1}} \frac{1 - f_{2hi}}{1 - m_{hi}^{-1}}}$ で、ブートストラップ標本に含め、(3) に進む。 SSU_{hij} をブートストラップ標本に含めない場合は、 PSU_{hi} の m_{hi} 個の SSU から、1 つの SSU をランダムに取り出しブートストラップ標本に含める。
- (3) (2) で SSU_{hij} をブートストラップ標本に含めた場合、 SSU_{hij} の k 番目の最終抽出単位 USU_{hijk} を確率 $r_{hij} = \sqrt{1 - \frac{f_{1h}}{p_h^{-1}} \frac{f_{2hi}}{q_{hi}^{-1}} \frac{1 - f_{3hij}}{1 - l_{hij}^{-1}}}$ でブートストラップ標本に含める。 USU_{hijk} をブートストラップ標本に含めない場合、 SSU_{hij} の l_{hij} 個の USU から、1 つの USU をランダムに取り出しブートストラップ標本に含める。

(1)~(3) を各層で行い、ブートストラップ標本を作成し、分位点推定値を計算する。これを B 回 (200 回から 500 回程度で十分である) 繰り返し、 B 個の分位点推定値の分散を計算することにより、BBE によるブートストラップ分散推定値が得られる。ただし、 p_h , q_{hi} , r_{hij} が 0 以上、1 以下でない場合、Short cut BBE は、実行不可能である。この場合は、General BBE が利用可能である (本稿で扱う問題に General BBE を利用する必要はない)。詳細は Funaoka, et al. (2006) を参照されたい。

参考文献

- [1] Funaoka, F., Saigo, H., Sitter, R.R., and Toida, T. (2006), Bernoulli Bootstrap for Stratified Multistage Sampling, *Survey Methodology*, 32, 151-156.
- [2] Shao, J., and Tu, D. (1995), *The Jackknife and Bootstrap*, Springer-Verlag.
- [3] 総務庁統計局 (2000) 『平成9年全国物価統計調査報告 (大規模店舗編, 小規模店舗編, 特売価格編, 日本の物価構造 (解説編))』, 日本統計協会.
- [4] 総務省統計局 (2005) 『平成14年全国物価統計調査報告 (大規模店舗編, 小規模店舗編, 特売価格編, 日本の物価構造 (解説編))』, 日本統計協会.
- [5] 菅 幹雄 (2004) 『物価指数の測定論』, 日本評論社.
- [6] 樋田 勉 (2006) 「平成9年全国物価統計調査の価格分布についての検討」, 『日本統計学会誌』, 35, 143-164.

(原稿提出日 平成19年9月12日)
(修正原稿提出日 平成19年11月7日)