

2019年全国家計構造調査 貯蓄額等における外れ値への 対処方法について

令和 2 年10月21日

総務省統計局

貯蓄額の集計における外れ値の考え方

貯蓄階級別の世帯数分布を見ると、高額データの多く存在

2014年全国消費実態調査 貯蓄額階級別世帯数分布（総世帯・全国・10万分比・不詳除く）

150万円未満	150万～300万	300万～450万	450万～600万	600万～750万	750万～900万	900万～1200万	1200万～1500万	1500万～2000万	2000万～3000万	3000万～4000万	4000万以上
13502	8314	7976	6770	5600	4538	7707	5774	6921	8474	4976	8235

4000万～5000万	5000万～6000万	6000万～7000万	7000万～8000万	8000万～9000万	9000万～1億	1億以上
2902	1709	1148	705	506	293	972

※4000万円以上の内訳は特別集計による

- 貯蓄額の分布を見る上で、高額データの存在は重要
 - 特に都道府県別の集計をする上で、高額データは平均値の推定誤差を過大にする要因となる可能性
- ➡ 都道府県別の平均値の表章が重要な調査であることを考慮し、貯蓄額が高額で出現頻度が極めて小さい世帯に対して付与するウエイトをごく小さな値とする（以下、「特別処理」という。）ことにより、所得資産集計体系の平均値に与える影響を最小化することを考える。
- ※上記の集計上の特別処理の導入に合わせて、都道府県別に特別処理の対象となった世帯数等を併せて公表することを検討

貯蓄額の外れ値への対処方針案

1. 貯蓄額に上限を設定する方法

- 世帯の貯蓄額に上限を設定し、上限を超えた世帯は特別処理の対象とする
- 上限値は調査データや過去の結果等を参考に決定
(2014年データの最大値は7億円弱)

※各都道府県で貯蓄額の水準は大きく異なるため、都道府県別に上限を決める必要
➡ 調査データを見ながらの作業となり、統一的な基準を作るのは困難

2. 平均値への「影響率」に上限を設定する方法

- 都道府県別の貯蓄額平均値に対して、個々の世帯のデータが与える「影響率」に上限を設定し、上限を超えた世帯は特別処理の対象とする
- 「影響率」の定義：

$$\text{影響率 (\%)} = \frac{\text{世帯の貯蓄額} \times \text{世帯のウエイト}}{\sum_{\text{各都道府県内の全世帯}} \text{世帯の貯蓄額} \times \text{世帯のウエイト}} \times 100$$

- 上限値は目標とする標準誤差率等を参考に決定

※各都道府県における貯蓄額の水準の違いに依存しない基準を設定可能
基準を低く設定しすぎると、出現頻度の小さくない世帯も処理対象となる可能性

都道府県別にみた影響率の大きい世帯数の分布

※2014年全国消費実態調査結果を特別集計。影響率の算出には「貯蓄・負債に関する結果」の集計用乗率を使用

	集計対象 世帯数	都道府県別平均値への影響率が基準を超える世帯の数				
		5%超	4%超	3%超	2%超	1%超
北海道	2099	0	0	1	1	5
青森	689	1	3	4	6	12
岩手	743	1	1	1	1	8
宮城	733	0	0	0	0	11
秋田	729	2	2	2	3	9
山形	710	0	1	1	3	10
福島	867	0	1	2	4	7
茨城	1369	0	0	0	0	3
栃木	818	0	0	0	0	4
群馬	854	0	0	0	1	3
埼玉	2705	0	0	0	0	1
千葉	2617	0	0	0	2	4
東京	2349	0	0	1	2	12
神奈川	2491	0	0	1	2	3
新潟	938	0	0	0	1	7
富山	719	0	0	0	1	6
石川	710	1	1	1	3	10
福井	720	0	0	0	1	4
山梨	700	0	0	0	1	7
長野	964	0	0	0	0	4
岐阜	905	0	0	0	1	8
静岡	1517	0	0	0	2	4
愛知	2396	0	0	0	0	4
三重	887	0	0	0	0	8

	集計対象 世帯数	都道府県別平均値への影響率が基準を超える世帯の数				
		5%超	4%超	3%超	2%超	1%超
滋賀	717	0	0	0	0	5
京都	760	2	2	2	4	9
大阪	2637	0	1	2	2	5
兵庫	2046	0	0	0	0	2
奈良	694	0	0	0	0	8
和歌山	695	0	0	0	3	12
鳥取	711	1	1	1	1	9
島根	717	0	1	1	1	8
岡山	761	0	1	1	2	7
広島	1024	0	0	1	2	9
山口	766	0	0	1	2	9
徳島	707	0	0	2	3	9
香川	725	0	0	0	1	6
愛媛	719	0	0	0	0	9
高知	712	2	2	2	4	13
福岡	1556	0	0	0	1	9
佐賀	726	0	0	0	2	13
長崎	689	0	1	2	3	8
熊本	694	1	1	2	3	9
大分	729	0	0	0	2	13
宮崎	725	1	1	2	3	12
鹿児島	734	0	0	0	3	12
沖縄	712	1	1	2	4	17
全国計	51185	13	21	35	81	367

基準を「3%超」としても半数以上の都道府県で基準を超える世帯が存在

➡ 出現がまれと見なせるのは「5%超」程度か