



総統調第 221 号

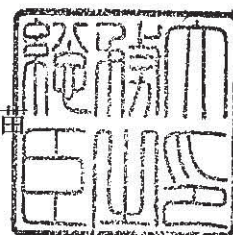
平成 26 年 9 月 10 日

統計委員会委員長

西村 清彦 殿

総務大臣

山本 早苗



諮問第 72 号

社会生活基本調査（調査票 B）に係る匿名データの作成  
について（諮問）

標記について、別紙のとおり作成するに当たり、統計法（平成 19 年法律第 53 号）  
第 35 条第 2 項の規定に基づき、統計委員会の意見を求める。

## 諮 問 の 概 要

社会生活基本調査（調査票B）（平成13年及び18年）について、統計法（平成19年法律第53号）第35条第1項の規定に基づき、以下のとおり匿名データの作成を行う予定であることから、同法同条第2項に基づき統計委員会に諮問を行うもの。

### 1 作成する匿名データの種類

平成13年及び18年社会生活基本調査（調査票B）の匿名データを作成する。

### 2 社会生活基本調査（調査票B）の匿名データを作成する理由

社会生活基本調査（調査票B）は、生活時間の配分や余暇時間における主な活動の状況など、国民の社会生活の実態を明らかにするため、世帯及びその世帯員を対象に平成13年から5年ごとに実施している統計調査であり、学術研究や高等教育において利用ニーズが高い調査であることから、匿名データを作成するものである。

### 3 匿名データの作成方法の概要

適用する匿名化措置は、以下のとおりである。

- ・元の統計調査のレコード全てを匿名データに用いるのではなく、それに間引きを施したものを用いる（レコードのリサンプリング）。
- ・直接的な識別情報は、レコードから全面的に削除する。また、レコードの配列順が意味を成さないように、無作為に並べ替えを行う（識別情報の削除等）。
- ・発生頻度の低いレコード又は特徴的な値があるレコードを含む世帯を削除する（裾切りによるレコード削除）。
- ・極端に大きな値は、上限値を設けて頭打ちにする（トップコーディング）。
- ・分類区分の程度は、詳細なものではなく、粗いものとする（リコーディング）。

## 社会生活基本調査 調査票 B（平成 13 年及び 18 年）の匿名データの作成方法（案）

社会生活基本調査については、「匿名データの作成・提供に係るガイドライン」（平成 21 年 2 月 17 日付け総務省政策統括官（統計基準担当）決定）に示された基準を踏まえ、調査票情報（個票データ）に対して以下の処理を施すことにより匿名データを作成する。

### 1 作成する調査年次

平成 13 年及び 18 年

### 2 ファイルの種類

世帯単位のファイル

※ 匿名データに収録される世帯については、特定の世帯員を除外する等の措置は行わず、全世帯員が収録される。

※ レコード編成としては、個人単位のレコード（2 日分の回答で 2 レコード）。なお、調査票 A と異なり生活行動のファイルは存在せず、生活時間のファイルのみ。

### 3 地域区分

全国

※ 社会生活基本調査の調査票 B は、標本サイズが約 1 万人（約 4 千世帯）となっており、社会生活基本調査 調査票 A の約 20 万人（約 8 万世帯）と比較して少数である。そのため、結果表章においても「全国」でしか表章を行っていない。

### 4 リサンプリング方法

リサンプリングは、世帯を単位としてまとめた上で、単純無作為抽出で世帯を抽出する方法により行いつつ、レコード（個人）を単位とするリサンプリング率が約 80%になるようにする。

### 5 情報の削除

#### (1) 直接的な識別情報の削除

調査地域を特定する調査区番号などの実査用の識別番号や調査客体を直接識別できる情報は削除するとともに、レコードは世帯を単位としてまとめた上で乱数により並び替える。

なお、電磁的記録である調査票情報には、調査票に文字で記入された「氏名」などは含まれていない。

「生年月」は、削除する（「年齢」に換算して利用）。

#### (2) 発生頻度の低い又は特徴的な値があるレコードの削除

① 世帯人員数9人以上の世帯

② 子供の数の多い世帯

国勢調査（母集団情報）において、親の年齢、子供の数及び住宅の所有の関係を組み合わせた結果、発生頻度の低い組合せに該当する世帯を削除する。

③ 母子世帯及び父子世帯

削除基準については、「②子供の数の多い世帯」と同様の基準とし、母又は父の年齢、子供の数及び住宅の所有の関係を組み合わせた結果、発生頻度の低い組合せに該当する世帯を削除する。

④ 三つ子以上がいる世帯

## 6 分類区分の再編

報告書の表章区分に合わせて分類区分を再編することを原則とする。また、報告書の表章区分が複数存在する場合には、より詳細な区分を使用する。

(1) 世帯員に関する項目

① 年齢

0～9歳は、各歳

10～84歳は、5歳階級

85歳以上は、トップコーディング

② 末子年齢

12歳以上トップコーディング

(0歳、1～2歳、3～5歳、6～8歳、9～11歳、12歳以上)

(2) 世帯に関する項目

特になし。

匿名データのチェックリスト（世帯調査用）  
「匿名データの作成・提供に係るガイドライン」  
（総務省政策統括官（統計基準担当）決定）準拠

匿名データを作成する統計データの名称及び年次

統計調査名：社会生活基本調査（調査票 B）  
調査年：平成 13 年及び 18 年

## 1 地理的情報

- (1) 提供するファイルにはどのレベルの地理的情報が含まれていますか。匿名化のために地理的情報を加工していますか。

①地理情報のレベル : 1 区分（全国）  
②地理情報の加工の有無： 有 無  
項目：市町村番号、調査区符号、世帯番号  
方法：削除

- (2) 直接的な地理的情報以外に地理的情報が明らかになるような情報がありますか。

地理情報以外の地理的情報： 有 無

- (3) 地域分析用に詳細な地理的情報を提供していますか。提供している場合、どのレベルの地理情報が含まれていますか。

地域分析用の地理情報提供の有無： 有 無  
地理情報のレベル：

- (4) ある特定の種類の施設であることが明らかになることはありますか。

特定の種類の施設の情報： 有 無

## 2 世帯の識別情報

- (1) 世帯の識別情報として考えられるデータ項目を挙げてください。

・世帯員数  
・世帯員の年齢構成  
・末子の年齢

- (2) それぞれの識別情報について、どのような匿名化措置を行っていますか。また、母集団に対する割合はどのようになっていますか。

匿名化措置の方法（項目ごとに記載してください。）  
項目：世帯員数  
方法：世帯人員が 9 人以上の世帯を削除

項目：世帯員の年齢構成

方法：同一年齢に3人（三つ子）以上がいる世帯を削除。また、調査対象となる世帯員の年齢を5歳階級にリコーディングし、85歳以上をトップコーディング

項目：末子の年齢

方法：12歳以上をトップコーディングし、6区分にリコーディング（0歳、1～2歳、3～5歳、6～8歳、9～11歳、12歳以上）

(3) 世帯単位のデータを提供することに対応して特別な匿名化措置を行っていますか。

特別な措置：行っている 行っていない

項目：子供の数の多い世帯

方法：国勢調査（母集団情報）において、親の年齢、子供の数及び住宅の所有の関係を組み合わせた結果、発生頻度の低い組合せに該当する世帯を削除

項目：母子世帯及び父子世帯

方法：子供の数の多い世帯と同様の基準とし、母又は父の年齢、子供の数及び住宅の所有の関係を組み合わせた結果、発生頻度の低い組合せに該当する世帯を削除

### 3 個人の識別情報

(1) 個人の識別情報として考えられるデータ項目を挙げてください。

・出生の年月

(2) それぞれの識別情報について、どのような匿名化措置を行っていますか。また、母集団に対する割合はどのようになっていますか。

匿名化措置の方法（項目ごとに記載してください。）

項目：出生の年月

方法：削除

### 4 誤差（ノイズ）

匿名化措置として、誤差を付加する方法を採っていますか。誤差を付加する方法を採っている場合には、その方法を記載してください。

誤差の付加：採用している 採用していない

方法：

### 5 リサンプリング

匿名化措置として、リサンプリングをしていますか。リサンプリングをしている場合には、その抽出方法と抽出率を記載してください。

リサンプリング：行っている 行っていない

方法：リサンプリング率が約80%となるように、世帯単位で単純無作為抽出を行う。

## 6 外部の情報

- (1) 個人・世帯を特定できる可能性のある外部の情報は存在しますか。

外部情報： 有 無

- (2) 母集団情報として利用している情報は何か。

国勢調査の調査区名簿

行政記録から作成した名簿（行政記録の名称： ）

その他（具体的に記載 ）

※母集団情報を取り扱う者の範囲等、特記する事項があれば記載してください。

調査区内の世帯名簿は調査の一環として作成され、調査関係者以外の者は見ることはできない。

## 7 その他

- (1) データの一連番号、データの並び順について、何らかの匿名化措置を施していますか。

実施している 実施していない

方法：データは世帯単位で乱数により並び替え

- (2) サンプル情報によって、地理的情報以外に特定の地域や集団であることが明らかになる可能性はありますか。

ない

- (3) 提供時期と調査時点とはどの程度の期間が開いていますか。

5年以上

- (4) そのほか、データを匿名化するに当たり、措置していることがありますか。

ない

掲載項目欄の凡例

- ：そのまま提供
- ：匿名化措置を講じて提供
- －：提供しない

識別情報欄の凡例

- レ：識別情報として考えられる項目

調査項目	提供項目	識別情報	匿名化措置	備考
調査員記入欄	調査区符号	－	レ	
	世帯番号	－	レ	
	世帯員番号	○		
	10歳以上の世帯員数	●	レ	世帯人員が9人以上の世帯を削除
	10歳未満の世帯員数	●	レ	
一人の世帯（単身赴任、その他）	○			
男女の別	○			
世帯主との続き柄	○			
出生の年月	－	レ		
年齢	●	レ	・5歳階級で提供。85歳以上をトップコーディング ・同一年齢に3人（三つ子）以上がいる世帯を削除	
配偶者の有無	○			
教育	○			
ふだん自分の用途で携帯電話やパソコンなどを使用していますか	○			
ふだん家族の介護をしていますか	○			
ふだん仕事をしていますか	○			
勤めか自営かの別	○			
本人の仕事の種類（職業符号）	○			
ふだんの1週間の就業時間	○			
生活時間について				
日にち	○			
この日は次のいずれの日でしたか	○			
この日の天気はどうでしたか	○			
世帯について	住居の種類	○		
	居住室数	○		
	自家用車の有無	○		
	世帯の年間収入	○		
	ふだん世帯以外の人から介護の手助けを受けていますか（介護支援の利用状況）	○		
不在者の有無	○			
10歳未満の人について				
世帯主との続き柄	○			
年齢	●	レ	・各歳で提供 ・同一年齢に3人（三つ子）以上がいる世帯を削除	
在学・在園の状況	○			
記入状況1日目か2日目か	○			
世帯の家族類型	○			
共働きか否か	○			
末子の年齢	●	レ	・12歳以上をトップコーディングし、6区分で提供（0歳、1～2歳、3～5歳、6～8歳、9～11歳、12歳以上）	
6歳未満の子供の有無・人数	○			
6歳未満の子供の保育の状況	○			
6～9歳の子供の保育の状況	○			
曜日	○			
その日の行動の種類	○			
人口乗率	○			
生活行動について（生活時間編）				
行動の種類（行動符号）	○		73区分で提供	
主行動				
行動時間	○			
インターネットの利用	○			
場所	○			
一緒にいた人	○			
同時行動				
行動時間	○			
インターネットの利用	○			
場所	○			
一緒にいた人	○			
生活時間について（時間帯編）				
行動時間	○		96区分で提供（15分単位：24時間×4）	
主行動				
行動の種類（行動符号）	○		73区分で提供	
インターネットの利用	○			
場所	○			
一緒にいた人	○			
同時行動				
行動の種類（行動符号）	○		73区分で提供	
インターネットの利用	○			
場所	○			
一緒にいた人	○			



社会生活基本調査（調査票B）匿名データの提供項目と匿名化措置一覧  
（平成18年）

掲載項目欄の凡例

- ：そのまま提供
- ：匿名化措置を講じて提供
- －：提供しない

識別情報欄の凡例

- レ：識別情報として考えられる項目

調査項目	提供項目	識別情報	匿名化措置	備考
調査区符号	－	レ		
調査世帯番号	－	レ		
世帯員番号	○			
10歳以上の世帯員数	●	レ	世帯人員が9人以上の世帯を削除	
10歳未満の世帯員数	●	レ		
一人の世帯（単身赴任、その他）	○			
男女の別	○			
世帯主との続き柄	○			
出生の年月	－	レ		
年齢	●	レ	・5歳階級で提供。85歳以上をトップコーディング ・同一年齢に3人（三つ子）以上がいる世帯を削除	
配偶者の有無	○			
教育	○			
ふだん自分の用途で携帯電話やパソコンなどを使用していますか	○			
ふだん家族の介護をしていますか	○			
ふだん仕事をしていますか	○			
勤めか自営かの別	○			
本人の仕事の種類（職業符号）	○			
ふだんの1週間の就業時間	○			
世帯について				
住居の種類	○			
居室数	○			
自家用車の有無	○			
世帯の年間収入	○			
ふだん世帯以外の人から介護の手助けを受けていますか（介護支援の利用状況）	○			
不在者の有無	○			
10歳未満の人について				
世帯主との続き柄	○			
年齢	●	レ	・各歳で提供 ・同一年齢に3人（三つ子）以上がいる世帯を削除	
在学・在園の状況	○			
集計用乗率	○			
ライフステージ	○		18区分で提供	
世帯の家族類型	○			
共働きか否か	○			
末子の年齢	●	レ	・12歳以上をトップコーディングし、6区分で提供（0歳、1～2歳、3～5歳、6～8歳、9～11歳、12歳以上）	
6歳未満の子供の有無・人数	○			
6歳未満の子供の保育の状況	○			
6～9歳の子供の保育の状況	○			
記入状況 1日目か2日目か	○			
日にち	○			
この日は次のいずれの日でしたか	○			
曜日	○			
その日の行動の種類	○			
この日の天気	○			
夫の家事時間階級	○			
妻の家事時間階級	○			
夫の育児時間階級	○			
妻の育児時間階級	○			
生活行動について（生活時間編）				
行動の種類（行動符号）	○		112区分で提供	
行動の有無	○			
インターネットの利用の有無	○			
場所	○			
一緒にいた人	○			
行動時間	○			
行動の種類（行動符号EU区分・A区分）	○		36区分で提供	
行動の有無	○			
行動時間	○			
同時行動（組み合わせの符号）	○		15区分で提供	
行動の有無	○			
行動時間	○			
生活時間について（時間帯編）				
行動時間	○		96区分で提供（15分単位：24時間×4）	
インターネットの利用の有無	○			
場所	○			
一緒にいた人	○			
主行動（行動符号）	○			
同時行動（行動符号）	○			

## 平成13年及び18年社会生活基本調査（調査票B）の概要

### 調査の目的

社会生活基本調査は、国民の生活時間の配分について調査し、国民の社会生活の実態を明らかにすることにより、各種行政施策の基礎資料を得ることを目的とし、昭和51年の第1回調査以来5年ごとに実施されている。

調査票B（アフターコード方式）については、平成13年より、従前の調査票である調査票A（プリコード方式）に加え、生活時間の配分の詳細な結果を得るために新たに導入した。

### 調査の時期

調査は、10月20日現在で行った。

ただし、「1日の生活時間」については、調査日の前後9日間のうち、調査区ごとに指定した連続する2日間について調査した。

### 調査の法的根拠

この調査は、統計法（昭和22年法律第18号）に基づく指定統計調査（指定統計第114号）で、社会生活基本調査規則（昭和56年総理府令第38号）に基づいて実施した。

### 調査の対象

国勢調査調査区から約350調査区を選定し、その調査区内から選定された約4千世帯に居住する10歳以上の世帯員約1万人を対象とした。

ただし、次の者は調査の対象から除いた。

ア 外国の外交団、領事団及び軍隊の構成員（家族、随員及び随員の家族を含む。）

イ 自衛隊の営舎内又は艦船内の居住者

ウ 刑務所・拘置所に収容されている者のうち刑の確定している者及び少年院・婦人補導院の在院者

エ 社会福祉施設の入所者

オ 病院、療養所等の入院患者

カ 水上に住居を有する者

### 調査の方法

この調査は、総務大臣（統計局長）－都道府県知事－統計調査員（指導員）－統計調査員（調査員）－調査世帯の流れにより行った。

調査は、都道府県知事により任命された統計調査員が調査票を世帯ごとに配布し、収集する方法により行った。また、調査票への記入は、原則として世帯が行った。

### 集計及び結果の公表

総務省統計局に提出された調査票は、独立行政法人統計センターにおいて集計される。

調査の結果は、インターネット等で公表する。

## 平成13年及び18年社会生活基本調査（調査票B）の標本設計の概要

**調査の範囲**

調査の範囲は、国勢調査調査区（以下「調査区」という。）のうち、以下の「除外する調査区」を除く調査区内の世帯の世帯人員である。

（除外する調査区）

- a 山岳・森林・原野地帯等の調査区
- b 大きな工場・学校等のある調査区
- c 社会施設・大きな病院のある調査区
- d 刑務所・拘置所のある調査区
- e 自衛隊地域の調査区
- f 駐留軍地域の調査区
- g 水面調査区

**標本抽出方法**

第1次抽出単位を調査区とし、第2次抽出単位を世帯とする層化2段抽出法である。

第1次抽出では、都道府県ごとに、市区町村コード、調査区番号順に配列し、この配列を基に全調査区の人口を累積し、累積した人口に対して確率比例系統抽出により、調査区を抽出した。

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世帯を抽出した。

**推定方法**

調査結果のうち、「行動者数」及び10歳以上人口は、別途推計した男女、年齢階級別人口を基準人口とする比推定によった。

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（又は「行動者数」）の推定値で除して求めた。

「行動者率」は、推定値の百分比として算出した。

**推定値の標本誤差**

標本誤差の算出は、副標本方式により行った。

副標本方式による標本誤差の計算は、事後的に4組の副標本を設定し、この4組の副標本ごとに算出された推定値を用いて算出した。

## (1) 情報の削除

ア：レコードのリサンプリング	元の統計調査のレコード全てを匿名データに用いるのではなく、それに間引きを施したものをを用いる。
イ：識別情報の削除等	識別情報は、レコードから全面的に削除する。 また、レコードの配列順が意味を成さないように、無作為に並べ替えを行う。
ウ：裾切りによるレコード削除	特徴的な識別情報の値があるレコードは、削除する。

## (2) 識別情報の階級区分統合

ア：トップコーディング	極端に大きな値は、上限値を設けて頭打ちにする。
イ：リコーディング	分類事項の程度は、詳細なものではなく、粗いものとする。

