歳時記

白鳥

冬の訪れとともにオホーツクの海をこえた白鳥は，再び，その優美な姿を私たちに見せてくれた。千里の道を遠しとせず，我が家を求める鳥の「渡り」に，私たちは神秘感さえ覚える。その手がかりは，鳥の体内時計と太陽コンパスのかみ合わせにあると聞く。

白鳥は，洋の東西を問わず，文学の素材にもちいられてきた。『古事記』のなかにも，東征の帰途，悲壮な最後をとげた倭建命（ヤマトタケルノミコト）が白鳥となって天翔けていく光景がある。きわめてロマンチックな想いをいだかせるが，それは感傷であって，眼前の白鳥は，生きとし生けるもののたくましさにあふれている。

白鳥は，体内の代謝が旺盛で，栄養の消費もはげしい。生物学者によれば，「渡り」という生活法を獲得したからこそ，今日まで繁栄しつづけてきたのだと言う。春の「渡り」にさきがけて，今，白鳥は体内に脂肪を蓄えている。そのうるわしき姿だけをいつくしむのは，私たちの身勝手であるらしい。

２月のおもな行事

- ○１日　1980年世界農林業センサス調査日
- ○５～６日　毎月勤労統計調査ブロック会議（東京都）
  個人企業経済調査地方別事務打合会議（長野県）
- ○７～８日　関東ブロック統計主管課長会議（神奈川県）
  関東ブロック県民所得研究会（山梨県）
  農林業センサス関東ブロック会議（大洗町）
- ○11日　建国記念日
- ○12日　全国統計大会班長会議
- ○13～14日　文部省関係説明会（東京都）
- ○14～15日　家計調査特別講習会（東京都）
- ○25～27日　常住人口調査事務打合会

さいじき

●シリーズ「統計研修から」

統計データの見方・表わし方(4)

――比率の種類と使い方２――

1. 意味上の区分 ――フローとストック――

比率を議論するとき，前回は，分母・分子のとり方に関連して比率の種類を区分してきました。今回は，そのテクニックの面での区分ではなく，意味の上での区分を強調したいと思います。ここで取りあげたいのは，フローとストックの概念の違いです。

フローとストックという概念は，いろいろな分野で使われていますが，これは統計の基礎概念でもあるのです。例えば，貯金を頭においてください。今月いくら貯金したかということと，月末現在いくら貯金をもっているかということでは概念が違います。前者のように，ある期間における現象の発生率を表わすものがフローです。そして後者のように，その結果としてきまる水準を表わすものがストックです。この区分は貯金だけに限らず，あらゆる分野で重要なものです。(表一１参照。)

表一１　フローとストックの例

	フロー	ストック
例	新規預入	預金現在
	出生数(死亡数)	人口数
	罹患率	有病率

大事なのは，フローはある期間に対応する情報であり，ストックはある時点に対応する情報であるということです。これが基本的なポイントです。フローとストックを混ぜて使うと，いろいろな誤った見方がでてきます。その扱い方をわきまえていなければなりません。そこで，フローとストックの間違いやすい事例として，〔例一１〕～〔例一３〕をあげておきます。

2. 比率の使い方 ――その１――

〔例一１〕をみてください。

いわゆる「罹病率」は非常に曖昧な言葉です。この分野の専門家は罹病率と言わないで，発生率とか有病率という言葉を使い分けています。発生率というのはフローの概念で，１年なら１年の期間にどれだけ病人が発生したかを表わしています。それに対して，有病率というのはストックの概念で，ある特定の時点で病気にかかっている人の数を表わしています。当然，その調査方法も異なります。発生率を調べるためには，１年の間に新しく病気にかかった人を何らかの方法で調べなければなりません。有病率を調べるためには，ある時点(何月何日現在)を固定して病気にかかっている人を調べていくことになるわけです。２通りの系統の数字はハッキリ違います。発生率から言えば，風邪のような呼吸器系疾患あるいは消化器系疾患というのが多く，心臓病のような循環器系疾患は少ないのです。そして有病率から言えば，風邪は比較的少なく心臓病みたいなものが多くなります。循環器系疾患の場合は一度病気にかかるとなかなか治りにくいからです。従って，発生率と有病率を混合するととんでもない間違った結論になってしまいます。

〔例一１〕 罹病率の表わし方について次の表に示す２通り(発生率・有病率)の方法がある。その差を説明せよ。

発生率と有病率

区分	発生率 (人口千人当たり年間罹患者数)	有病率 (人口千人当たり罹患者数)	罹患1回当たり罹病回数
全傷病	2,394.1	63.6	14.4
循環器系疾患	53.9	10.6	82.4
消化器系疾患	449.1	14.3	6.5
呼吸器系疾患	952.6	7.0	17.2

3. 比率の使い方 ――その２――

〔例一２〕をみてください。

〔例一２〕 心臓血管系疾患と消化器系疾患について，それぞれの死因による死亡者数(人口100万人当たり)とそれぞれの病気に罹病しているものの数(人口千人当たり)を調べたところ次のようになっていました。

	死亡者数 100万人につき	罹病者数 千人につき	致死率 千人につき
心臓血管系疾患	260	10	26
消化器系疾患	12	12	1

ある人が，これをつかって，病気の致死率すなわち罹病している人が死亡に至る確率を上表の３番目の欄のように計算していました。この計算方法は妥当ですか。

総理府統計局参事官　上田尚一

かんたんな計算がなされています。計算ちがいはないようですが，どこがおかしいのでしょうか。致死率とは，病気が死亡につながるかどうかという危険性を表わしたものです。そう考えれば，致死率を計算するため，病人の数を分母にとって，死亡者の数を分子にとるというのは分かります。例えば，心臓血管系疾患の致死率26（千人につき）というのは，

$$\frac{260}{1000000} \Big/ \frac{10}{1000} = 0.026$$

という計算をしているわけです。計算自体には誤りはありませんが，もっと根本的なところに誤りがあります。フローとストックの概念を思い出してください。例題の3行目に「それぞれの病気に罹病しているものの数」という表現があります。ある時現在に病気の状態にある人を取りあげているわけですから，これはストックです。それに対して，死亡者というのは発生ですからフローです。そうすると，期間に対応するわけです。ところが，死亡者数260人という数字には単位がありません。年間の数なのか月間の数なのか，単位を忘れているのです。先程の計算に単位を付け加えれば，

$$\frac{260\text{人/年}}{1000000\text{人}} \Big/ \frac{10\text{人}}{1000\text{人}} = 0.026\text{人/年}$$

という式になります。この数字は，1年間に1000人当たり26人死んでいくということです。ですから，例題の主旨である致死率（病気にかかった人がどのくらい死ねか）と無関係ではありませんが，1年当たりという条件が入りこんでおり，解釈しにくい数字になっています。この計算ででてきた数字は，その病気にかかって治るまでの期間が1年あるとすれば，その間に何人の割合で死ぬかという数字です。正しい意味で致死率を出そうとすると，この数字を修正しなければなりません。この病気にかかって治るまでに3年かかるとすれば，致死率はこの数字のx倍（3倍とは限りません）になるということです。

こうした誤解をさけるためには，分母・分子ともにフローの数字にすれば何ということもありません。分母にはある期間中に病気にかかった人の数を，分子にはそれと同じ期間に死んだ人の数をとるわけです。基本的には，フローとストックの数字を混ぜた比率を作らないことです。しかし，実際問題としては，混ぜた比率を使わなければならない場合が2つあります。その1つは，〔例―2〕の病人の数のように統計がとりにくい場合です。本来はフローに対するフローの比率にしたいが，分母の病人の数がつかみにくいので，ある時点の病人の数を分母にとり，後でややこしい修正をするわけです。もう1つは，意識的にストックに対するフローの比率を使うことの必然的な理由がある場合です。この場合については，4で述べます。

4．比率の作り方 ――因果関係――

〔例―3〕 次の表の見方を考えよ。どんな指標をつくって，どう対比するのか。

人種別死亡原因ー年齢50～59歳の米国人の死亡者について

集団区分	標識区分		
	ガン	その他	計
白人	140	1,055	1,195
黒人	9	170	179
計	149	1,125	1,374

（単位：1,000人）

ここでのねらいは，因果関係を議論するということです。因果関係をみるときには，ストックに対するフローの比率を使うことがよく行なわれます。フローとストックはお互いに関連性をもちながら変化していくわけですから，その関連の度合をみていくためにフローとストックを混ぜて使うのです。

ある問題意識があって，こういう表が作られています。その問題意識に応じた統計の使い方・表わし方を考えなければいけません。この問題の背後には，ガンにかかる割合（あるいはガンで死亡する割合）が人種によって違うのではないかという問題意識があるのです。人種の違いを原因に想定して，それがガンの発生という結果に対して，どの程度の差をもたらしているかをみるのがこの表の目的です。例えば，白人のガンの発生率（この場合は病気の発生ではなく，死亡者の中のガンによる死亡）については，$\frac{140}{1195}$という比率を作ればいいわけです。因果関係をみるために，しばしば統計が使われます。そういう場合，何を原因とみなし何を結果とみなして統計数字をみようとしているのかを想定し，因に相当するものを分母に，果に相当するものを分子にとります。これが比率を作る場合の基本的な見方です。このことを前提にして，〔例―4〕に入ります。

〔例―４〕　２つの都市Ａ・Ｂについて，カラーテレビの普及率の変せんを調べたところ，次の結果が得られました。

	43 年	44 年	45 年
Ａ　市	60	70	80
Ｂ　市	20	30	40

数字は100世帯当たりのカラーテレビ所有世帯数です。

ある人がこれにもとづいて「44年から45年にかけての普及率の伸びは，Ａ市よりＢ市のほうがいちじるしかった」と説明しています。この説明は適当といえますか。

カラーテレビの普及率の伸びという問題意識をもって，この統計を使いたい。その為には，この数字を生で見るのではなく，何らかの加工をしなければいけません。因果関係をどう考えるかが，この問題に対するカギになっています。前回に習った変化率$\left(\frac{当年の値}{前年の値}\right)$によるのが一応の常識です。しかし，実際の問題を扱うときは必ずしもそうするとは限りません。これがヒントです。

この例題については，３通りの見方が考えられます。その１つは，表―２のような見方です。

表―２　伸び率の見方 (1)

	44 年	45 年	その伸び
Ａ　市	70	80	10
Ｂ　市	30	40	10

⇨ 伸びが等しい

この数字は，44年から45年にかけて，１年間にＡ市・Ｂ市それぞれの市の10％の世帯がテレビを買ったことを表わしています。既に買って持っていた世帯は度外視して，新しく買った世帯がテレビ普及率の数字を「プラス10％」した――そういう意味で，両市とも同じだったわけです。逆にいうと，44年から45年の期間における普及率の差に注目して，伸びが等しいと見ているわけです。これも１つの見方です。しかし，何となく物足りない気がしませんか。そもそも，70から80に伸びることと30から40に伸びることとは差を計算すれば同じ10の伸びですが，それが同じ意味をもつだろうかということです。そこで，２番目の見方として表―３のような変化率の考えがでてきます。

表―３　伸び率の見方 (2)

	44 年	45 年	その伸び
Ａ　市	70	80	14％
Ｂ　市	30	40	33％

⇨ Ｂ市の方が伸びた

例えば，Ａ市では70から80に伸びた。変化率を用いると$\frac{80}{70}=1.14$ですから，14％伸びたということです。これが普通の考え方のように見えます。しかし，これについても異論はないでしょうか。極端な言い方をすれば，０から10へ伸びたとき伸び率は無限大でしょうか。また，10から20へ伸びたとき100％増加したわけですが，この100％に意味があるでしょうか。変化率を出すときには，前年の値に対する当年の値という使い方をしています。何％から何％へ変わったんだという結果をそのまま「記述する」のでしたら，これでいいのです。しかし「原因――結果」ということを考えに入れると，話はもっと突込まなければなりません。そこで，３番目の見方として表―４があります。

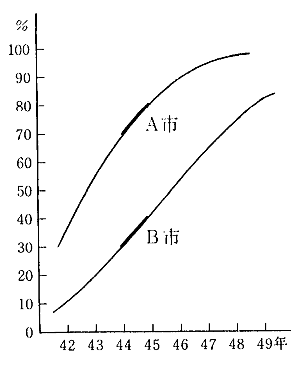
表―４　伸び率の見方 (3)

	44 年	45 年	その伸び
Ａ　市	70	80	33％
Ｂ　市	30	40	14％

⇨ Ａ市の方が伸びた

どうして上のような結果になったか分かりますか。70から80に伸びたことに対する評価，30から40に伸びたことに対する評価が違うことは感じで分かると思います。30から40へ伸ばすことは簡単ですが，70から80へ伸ばすことは容易ではありません。数字のうえでは同じ10の伸びであっても，伸び方の容易さから言うと全く違います。こういう現象を議論するとき，比率を出すことの背後には隠れた仮定――いくらでも数字は伸びるという仮定――があるわけです。ところが，数字の値に限界がある場合は，変化率を常識通りに出していいのかどうか考えなければいけません。

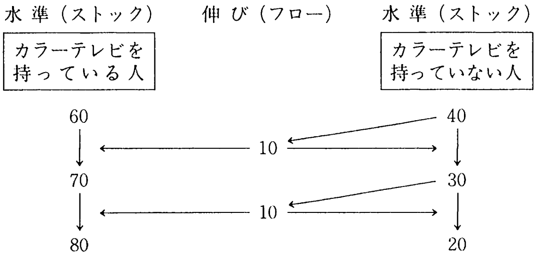
図―１　伸び率の予測



有界なデータについて変化率をつくる場合，変化率＝期間中の変化／期首の値　としてよいとは限らないことを強調する意味で，有界変化率という言葉をつかうことがあります。この考え方の背後には，100％という限界があって，限界に近づけば伸び率が図―1のように下ってくるという予測があります。A市・B市とも普通はそれぞれ実線のように予測されるのに対して言えば，A市は傾きが大きく，B市はゆるやかになっていると言えるわけです。ですから，限界がある場合には，普通の変化率の出し方には問題があります。普通の変化率の出し方が万能とは思わないでください。とくに限界があると考えられるデータについては注意が必要です。言わば，「同一条件下で予想される傾向線を描いて，それとの対比で問題を考えよ」というのが理論的な答え方ですが，この例に限ると，次のように考えればよいのです。

答を出しましょう。A市は$\frac{10}{100-70}=33\%$，B市は$\frac{10}{100-30}=14\%$として計算したわけです。なぜそうしたのか，図―2を見てください。

図―2　カラーテレビのフローとストック（A市の場合）



データの見方として，図の左側の

$$60 \xrightarrow{\overset{10}{\downarrow}} 70 \xrightarrow{\overset{10}{\downarrow}} 80$$

という見方があります。これは必ずしも因果関係ではなく，むしろ会計計算です。ストックがあって，その変化としてフローがあるということです。これは先程の2番目の出し方ですが，この考え方は因果関係をみるというよりも，ストックに対して新しくフローが入ってきたためストックが大きくなった事実を，言わば会計計算的に追いかけているわけです。言わば，データを記述する観点にたっています。しかし，テレビを買うという現象がどういう形で起こるのかに着目すれば，図の右側の見方をしたくなってきます。テレビを購入するという行為は，テレビを持たない人の側から発生しているからです。つまり，購入という行為は，40のうちから10，30のうちから10発生したんだという見方をします。どこから発生したのかその原因を想定して，その原因をもつような世帯の数を分母にとるということです。そう考えれば，A市$\frac{10}{100-70}$，B市$\frac{10}{100-30}$という計算になります。

2番目，3番目のいずれの考え方をとるかは，要するに数字の見方の問題です。見る立場如何でどれをとるかがきまってくるのです。「発生の原因を頭において，因に相当するものを分母にとり結果を分子にとる見方」にたつとすれば，「発生がどの範囲から生ずるか」を考えた上，この例では$\frac{\text{購入者}}{\text{持っていない人}}$の形で変化率を出すことになるわけです。

以上がフローとストックの扱いに関連した比率のとり方ですが，さらに1つつけ加えておくべきことがあります。たとえば，毎月1件づつ（たとえば病人が）発生して3ヵ月間状態（たとえば病気中）がつづく場合，ストックの数字（病人の現在数）は3です。同じく毎月1件づつ発生しても，状態の継続期間が6ヵ月になったとすれば，ストックの数字は6となります。フローの意味では同じでも，ストックでみるとちがうのです。このように，ストックの数字は，発生率だけでなく継続期間にも関係をもつ量です。だから，病人の数をフローでみるのか，ストックでみるのかが問題とされるのであり，失業者の数をフローでみるのと，ストックでみるのとはちがったことになるのです。たとえば，「失業給付金の給付期間を長くすると失業者がふえる」ということになるわけですから，雇用市場の影響か，保険制度の影響かが判断しにくくなることがあり得ます。このように，ストックの数字の解釈には，期間に関連した側面を含んでいることからくるむづかしさがあります。

統計数字の見方を議論するとき，世の中の制度はどうなっているんだということを，いつも頭の中に入れておかないと，統計数字を正しく読むことはできません。

編集子より；このシリーズは，上田先生が昭和54年3月に総理府統計研修所で講義されたものを収録・編集したものです。

◇統計の窓

統　計　と　数　学……………………

私はよく学生に，「近くて遠きは統計と数学」とか「統計と数学は似て非なるもの」という話をする。あるいは多少逆説的な言い方としては「あまり数学に強くなると統計のジャマになる」ともいう。

統計が扱う数字と，数学が扱う数字とが全くの異物であることを常に念頭におくことが如何に大切であるかは，これを何度くり返し力説しても決して十分ではないと私は考えている。

なぜこうしたことを改めて言う必要があるのかと思う方は本稿を読まれる必要はない。しかし，統計を応用数学の一部だと考えている方には是非御一読願いたいし，また，本誌の読者の大多数は統計マンであろうが，その方々の日常の仕事上の悩みの多くが数学とは全く関係ないこと，を改めて再確認していただきたいのである。換言すれば統計の特殊性を再認識して社会経済現象の数量的表現(＝統計)の有用性の限界を知っていただきたいのである。統計はたしかに有用ではあるが万能ではない。

しからば統計と数学，正確には統計学と数学とはどこが異るのか。それは次の二点である。第一は前者が具体数のみを扱うのに対して後者は抽象数のみを扱う。第二は前者が社会的集団のみをあらわすのに対して後者は抽象数のすべて(数量と空間)の関係を扱う。それでは具体数とは何か。それは歴史的数量といってもよい。つまり，特定の時と特定の場所でのみ存在する数量である。(たとえば1970年10月1日午前0時における日本国領土に常住する人口)。時と場所を特定しない統計(少くとも社会経済に関する統計)はあり得ない。統計はつねに時と場所を特定して実在する数量だからこれを歴史的数量と前記したのである。換言すれば再現不可能な数量なのである。よく，歴史はくり返す，というけれども真実の意味においては歴史的事実が再現されるということはない。つまり，くり返しのない一回性こそが具体数の特徴である。

それでは抽象数とは何か。それは時間と場所を特定されない数字，つまり人間の観念の世界だけに存在する数字である。たとえば1プラス1イコール2というときの1や2は実在物ではなく，一箇のリンゴや2箇のリンゴとは全く異った数字である。また，具体数と抽象数のちがいは，前者がいつでも個有の測定単位(たとえば，人，トン，箇，メートル，ドルなど)をもってあらわされるのに，後者はつねに無名数であるということもできよう。

つぎに社会的集団をあらわすのが統計である，という点についてのべよう。こうした言い方に対しては必ず二つの批判があらわれる。一つは自然測定値の統計はどうか，という批判。もう一つは社会的集団とはいえない歴史的数量，たとえば特定時点の日銀券発行残高などは立派な統計として通用しているが，それは日本銀行一行にかかわる数量であって社会的集団をあらわすとはいえない，という批判である。前者の批判(自然測定値)についていえば，たとえば物質の重量測定値はたしかに実在物の測定値であるが，測定は同一条件でくり返すことができ，したがって多数回の測定値の誤差分布が正規分布であることを前提して真値を算出しうることを忘れた批判ではないか。こうした測定値のあつまりを統計と呼称してはいけないとはいわないが，統計調査によって得られる統計とは全く異質のものである。統計調査は同一条件でのくり返しが全く不可能なのである。なぜならば調査客体の存在をめぐる社会的諸条件が時々刻々変動するからである。たとえば，国勢調査においてある地区で失敗があったとして，改めて後日に行う調査の結果は特定日時の国勢調査とはことなるものとなる。かって北海道のある町でこうした事実があった。統計局の国調報告書には後日行われたこの町の人口調査の結果と住民基本台帳からの推定人口数が記載されている。統計調査がもし自然測定のように同一条件でのくり返しが可能ならばこの北海道のある町の国調失敗事件は，国調やり直し，という単純な処理ですむことになる。もちろん，すべての自然測定がくり返し可能であるわけではなく，天体観測などでは対象物が自体変化するからそこには観測の一回性が認められる。しかしなおそれが統計調査と根本的に異るのは，天体観測は人間(観測者)と対象物との対応で測定値が得られるのに対し，統計調査では調査者も被調査者もともに人間であり，時には利害の一致しない人間の対応(意識の交流)の中からデータが生まれるという点にある。

さて，つぎに日銀券発行残高が統計であることは誰も異論はないものの，それが社会的集団をあらわしているかどうかという問題にうつろう。日銀券発行残高は日銀の帳簿に記録されている数字であるが，発行された日銀券そのものは企業や個人の所有物として存在しているので，日銀券発行残高の実体は社会的集団であると考えられる。

以上，統計と数学とのちがいを，具体数か抽象数か，社会的集団の反映かどうか，という二点にしぼってのべた。

統計の窓◇

筑波大学教授　三　潴　信　邦

ところで，統計もまた数字による表現であることは事実であるから，数学における加減乗除等々の計算手続を統計が援用することになる。しかし，計算技術を援用したからといって統計が抽象数に化けるわけではない。したがって，統計をたんなる数字としてではなく，社会的に生きている数字 ＝ 具体数 として扱うときに数学的技術を無条件で利用することがどんなに誤ったことであるかを十分に知ることが肝要ではないか。

科学における対象と方法，というやや抽象的ないい方をするならば，方法(数学的技術)は対象(統計)によってその用い方が限定されるのであってその逆ではない筈である。われわれが木材をノコギリで切り，紙はハサミで切り，牛肉はナイフで切るのはなぜだろう。木材はハサミでは切れないし，紙をノコギリで切らないのはなぜだろう。統計が主人で数学的技術はいわば従者であることを忘れてはいけない。

昭和55年国勢調査集計体系案　案内

集計区分		集計の性格	対象	公表の予定時期	表章地域	産業	職業	抽出方法	結果公表の方法	備考
区市町村別人口集計（全国、都道府県、市）	要計表による人口及び世帯数	世帯及び男女別人口の概数を早期に提供する	全数	昭和55年12月末	全国 都道府県 市区町村	—	—	—	新聞発表後報告書刊行，官報に公示	要計表を用いて集計する
	確定人口	人口及び世帯数の最終確定結果を提供する	全数	昭和56年10月末（目標）	全国 都道府県 市区町村	—	—	—	官報に公示	第1次基本集計結果による
抽出速報集計		基本集計，抽出詳細集計，従業地・通学地集計及び人口移動集計の一部を全国又は都道府県段階まで提供する	1　%	昭和56年3月末	全国 都道府県 人口50万以上の市	小分類	小分類	100分の1の世帯を抽出	報告書による	
第1次基本集計		人口，世帯及び住居に関する基本的な結果を市区町村段階まで提供する	全数	昭和56年10月末（目標）	全国 都道府県 市区町村	—	—	—	原則として報告書による	
第2次基本集計		産業，職業及び世帯の経済構成などに関する基本的な結果を市区町村段階まで提供する	全数	未定	全国 都道府県 市区町村	大分類	大分類	—	原則として報告書による	
調査区別集計		基本集計，従業地・通学地集計，並びに調査区特性に関する基本的な結果と標本調査用資料を調査区別に提供する	全数	未定	調査区	大分類	大分類	—	閲覧に供する	関連集計として地域メッシュ別集計を行う
従業地・通学地集計		従業地・通学地及び利用交通手段に関する基本的な結果を市区町村段階まで提供する	全数	未定	全国 都道府県 市区町村	大分類	大分類	—	原則として報告書による	
従業地・通学地集計		従業地・通学地及び利用交通手段に関する詳細な結果を一定規模以上の地域について提供する	20　%	未定	全国 都道府県 市区	中分類	中分類	5分の1の世帯を抽出	原則として報告書による	
人口移動集計		人口転出入状況に関する結果を市区町村段階まで提供する	全数	未定	全国 都道府県 市区町村	大分類	大分類	—	原則として報告書による	
抽出詳細集計		多重クロス表及び産業・職業などに関する詳細な結果を原則として都道府県段階まで提供する	20　%	未定	全国 都道府県 市区町村	小分類	小分類	5分の1の世帯を抽出	原則として報告書による	

1）調査区ごとに作成する。世帯名簿上の世帯数及び人口を市区町村，都道府県別に合算して作成するもの。
2）分割区のある調査区については分割区別に提供する。

（総理府統計局『統計ニュース第２号』から）