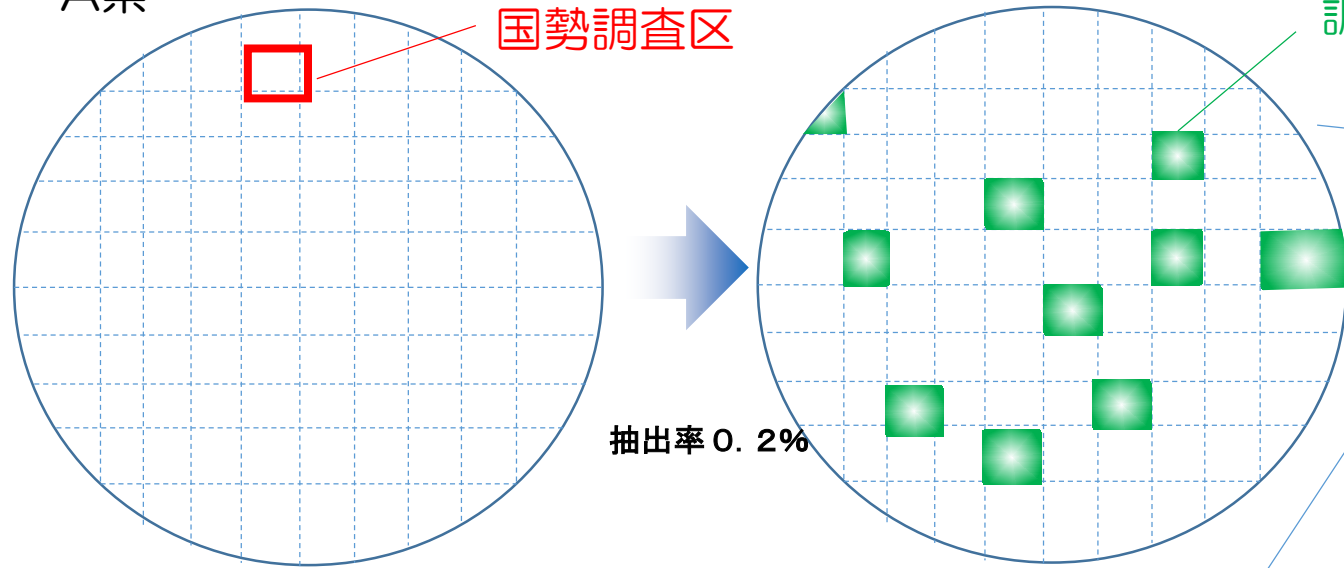


＜国民生活基礎調査は、**集落抽出法**を採用＞

- 地区全体が対象となるため、効率的な広報が可能（※「今年はこの地区が調査対象です」と事前に広報）
- 調査対象となる世帯が集中していることから、調査員の稼働効率が高く経費の抑制が可能
- 調査対象となる地区の全世帯が対象であるため、出現頻度の低い母子世帯等を漏れなく把握可能、また、調査対象者に対する調査協力の効率果的な説明が可能
- 多数の後続調査（出生動向基本調査等）が存在するため、後続調査とのマッチングが容易で効率的な調査が可能

A県



地域集計の精度確保のため、
県(政令市)毎に抽出地区数は同程度
(概ね100地区程度(政令市は60程度))

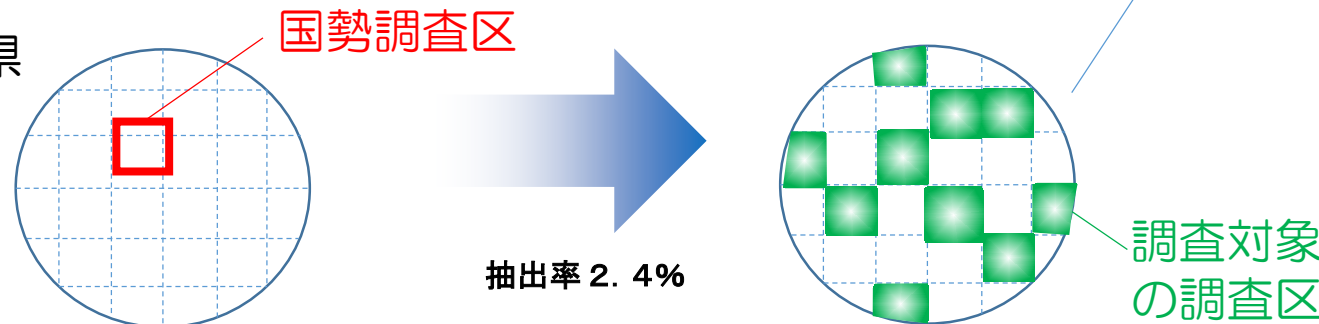


県(政令市)毎に抽出率が異なる
最大2.4%、最小0.2%程度

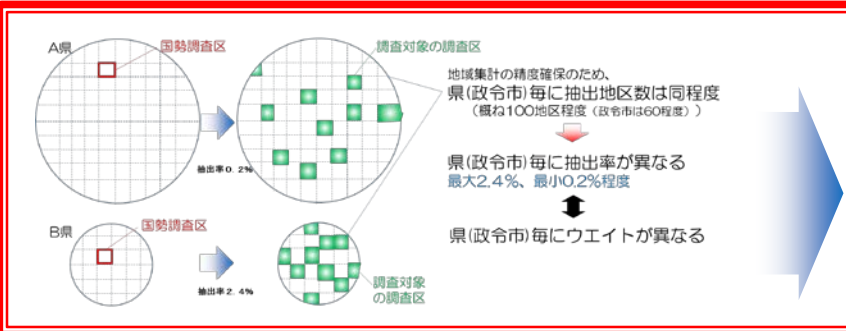


県(政令市)毎にウエイトが異なる

B県



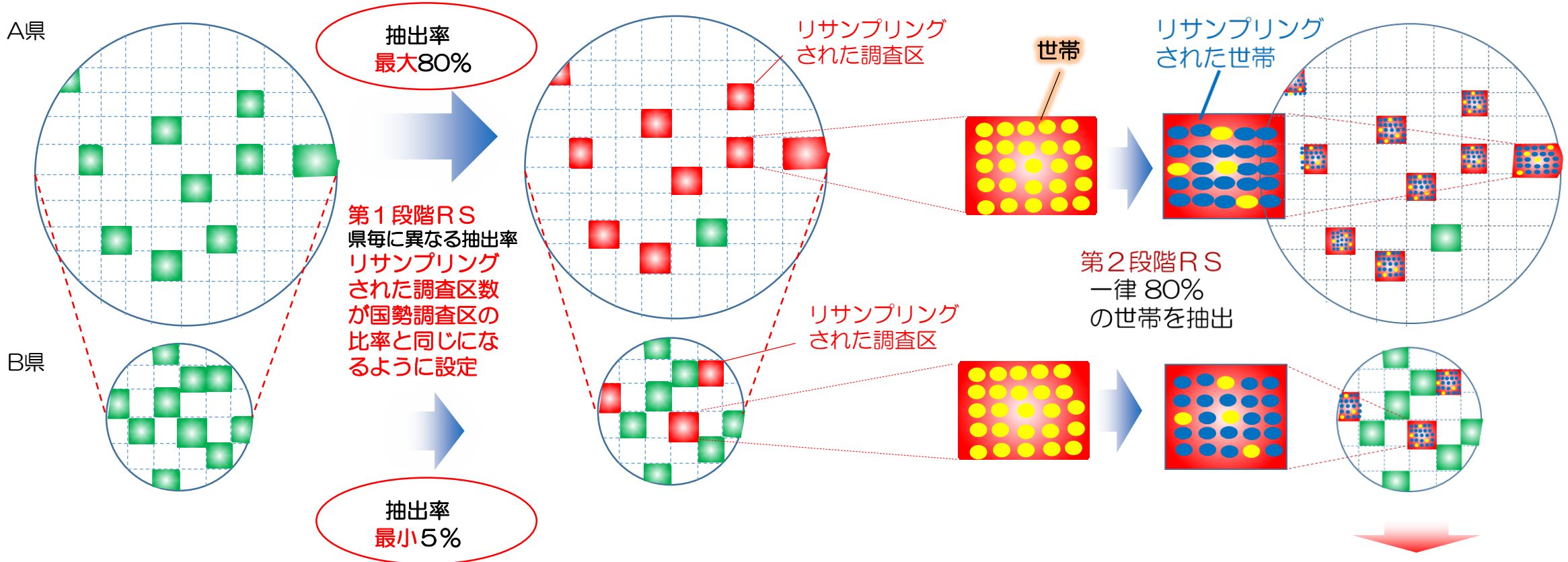
国民生活基礎調査の匿名データのリサンプリング



ウエイト情報 (=県特定情報) を消失させるリサンプリングが要件

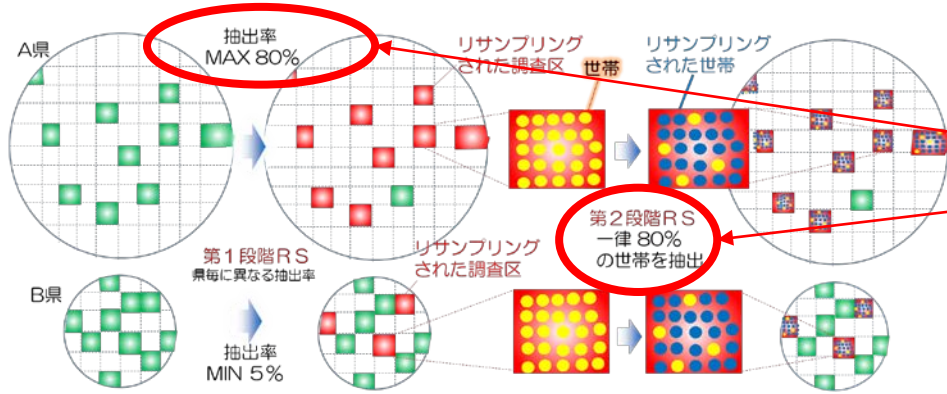
- 抽出率の逆数比で調査区をリサンプリング【第1段階RS(リサンプリング)】

※リサンプリングされた調査区内でも、世帯が特定しにくくするように、
調査区内の世帯をリサンプリング【第2段階RS(リサンプリング)】



●で抽出された世帯はいずれの県もウエイト(拡大乗数)は均等であり、単純な集計で全体推計が可能となる。

リサンプリング率について



リサンプリング率は80%値を使用

秘匿確率を求め、最適な割合を計算する時間・労力をかけるより、早く匿名データを作成し利用者に提供することを優先させ、先行事例である総務省採用の80%値を実績のある安全基準として使用

国民生活基礎調査の匿名データのリサンプリングの課題

これまでの考え方	ご意見等
<p>○匿名データで抽出された世帯の分布が（県別の世帯数の構成比）が、国勢調査の結果と同じであることを重視（中間年調査の標本設計を踏襲）</p> <p>○本体調査では、調査地区を広報しているためウエイト情報で県(政令市)から調査地区まで絞り込み、調査区から集落が特定できると回答者が探索される可能性があるので、集落が特定できない匿名化措置が必要</p> <p>○具体的には、以下の措置</p> <ul style="list-style-type: none"> ・地域に関する情報削除 ・リサンプリングによる拡大乗数の一律化 	<p>○世帯の分布（県別の世帯数の構成比）を国勢調査結果に合わせる必要性はないのではないか</p> <p>○本体調査の拡大乗数は、県レベルで一定なので、拡大乗数が付与されても、集落が特定されることはない</p> <p>○集落が判明しても、世帯（個人）が特定されることにはならないように、全世帯ではなく、8割の世帯を抽出しているのではないか</p> <p>○本体調査から世帯を直接抽出する方法がシンプル</p> <p>○都道府県レベルあるいは地域ブロックレベルであれば、匿名性は十分に確保できるのではないか</p> <p>○外観識別情報の組み合わせで標本一意にならなければ、匿名性は確保できるのではないか</p>



《今後の検討課題》

- ・匿名性確保について検討
- ・ダイレクト・リサンプリングの検討（公表値とのかい離分析等）