

あなたもいつか選ばれる？統計調査における標本とは

皆さん、国勢調査の回答にご協力いただきありがとうございました！

ところで、国勢調査では全ての人・世帯を対象として調査をしましたが、統計調査の中には対象となる集団の一部だけ(標本(ひょうほん))を調査するものもあります。今回は、この「標本」について説明します。

01 標本とはどのように選ばれるのか

□全数調査と標本調査□

「全数調査」は、ある性質を調べるために、調査の対象となる集団を全て調査しますが、「標本調査」は、統計的手法を用いて調査の対象となる集団の一部だけを調べ、その結果から全体を推計する調査方法です。

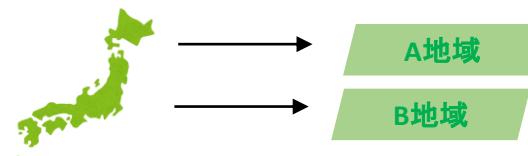


□標本抽出の方法□

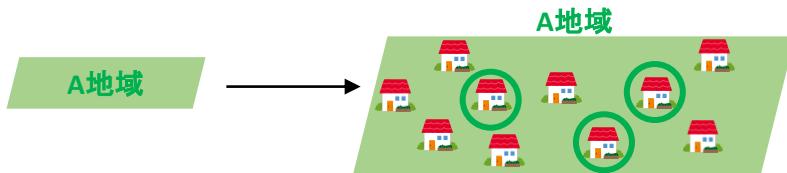
統計調査の標本を選ぶことを抽出といいます。今回はいくつかある抽出方法のうち、「層化2段抽出法(そうか2だんちゅうしゅつほう)」(例:労働力調査)を紹介します。

層化2段抽出法では、2段階の方法によって標本となる調査世帯が選ばれます。

【1段階目】まず、地域・産業のばらつきを考慮した上で、日本全体の縮図となるように地域を選ぶ。



【2段階目】つぎに、選ばれた地域にある世帯すべてを把握したうえで、無作為(誰でも同じ確率で選ばれる可能性がある方法)に標本となる世帯を選ぶ。



労働力調査では、毎月約4万世帯(就業状態を調査する15歳以上は約10万人)を対象に調査を実施し、1人分の回答は約1,100人の代表となっています！

02 標本誤差と必要な標本サイズ

標本調査は、全数調査に比べて手間や費用を省くことができる一方で、全数調査を行えば把握できるであろう“**真の値**”との誤差が発生することは避けられません。

この誤差を「**標本誤差**」といい、これをどの程度まで許容するのか、また、どの程度の「**信頼度**」(=確からしさ)」を求めるかによって、必要な「**標本サイズ**」(=標本として調査する数)」が決まります。

一般的に、標本誤差の許容範囲は3~5%、信頼度95%で標本サイズを計算する場合が多く、例えば調査対象者がある選択肢を回答した比率が50%である場合、標本誤差が5%、信頼度95%というのは、45%~55%の間に真の値がある確率が95%ということです。

なお、必要な標本サイズは下記の式で求めることができます。

(※最も簡単なケースであり、条件次第では複雑な計算が必要)

$$\text{標本サイズ} = \frac{\lambda^2 \times p(1-p)}{\text{標本誤差}^2}$$

λ : 信頼度から決まる値
※信頼度95%のとき λ = 1.96
p : 調査の結果得られる比率
※不明の場合標本サイズが最大となる p = 0.5

■ 標本サイズの計算(例) ■

県民の中から調査対象者を無作為に抽出し、アンケート調査(「はい」「いいえ」など二者択一で回答する形式)を実施したい。なお、標本誤差5%・信頼度95%に設定し、調査の結果得られる比率は不明

$$\frac{1.96^2 \times 0.5(1-0.5)}{0.05^2}$$



必要な標本サイズは**384**

03 標本に選ばれたら？

このように、統計調査の**標本**には誰もが選ばれる可能性があります。

標本に選ばれた世帯には、「**統計調査員**」が訪問し、調査の説明を行います。皆様の中には、もしかしたら「怪しい調査ではないか」など、不安になる方もおられるかもしれません。そこで、**調査にあたり皆様の疑問**にお答えします。



Q. 怪しい調査ではないか確認する方法はある？

A. 調査員は「**調査員証**」を携帯していますので、調査員証の提示を求めることで確認することができます。

Q. 調査員から、金銭の支払いや銀行口座の暗証番号を求められたんだけど…

A. 調査員は銀行口座の暗証番号を尋ねることはできません！それは、いわゆる「**かたり調査**（統計調査を装い個人情報等を詐取する行為）」である可能性が高いので、速やかに県統計調査課にお知らせください。

今回は統計調査の標本についてご紹介しました。標本に選ばれると「なぜ自分が？」と思われるかもしれません、標本に選ばれた際には、調査へのご協力をよろしくお願ひします。

【参考文献】総務省統計局「労働力調査」「なるほど統計学園」「Date StaRt」、(公財)統計情報研究開発センター「統計実務基礎知識」

