

# 調査における留意事項

広島大学 AI・データイノベーション教育研究センター

稲垣知宏

# 目標

大きな集団の特徴を調べる上で留意すべきことを理解し、調査方法の妥当性を検討できるようになる

この授業で紹介すること

- 標本調査における留意事項
- アンケート調査における留意事項

キーワード

母集団と標本抽出、無作為抽出、標本の大きさ、アンケート調査

# こんなことはありませんか？

1週間の学習時間について、大学生を対象に調査します。

どのようにすれば、信頼性のあるデータを集めることができるでしょうか。



A checklist form with four items, each with a checkbox and a text area. The first, third, and fourth items are checked.

<input checked="" type="checkbox"/>	_____
<input type="checkbox"/>	_____
<input checked="" type="checkbox"/>	_____
<input checked="" type="checkbox"/>	_____

# 集団の特徴を調べる

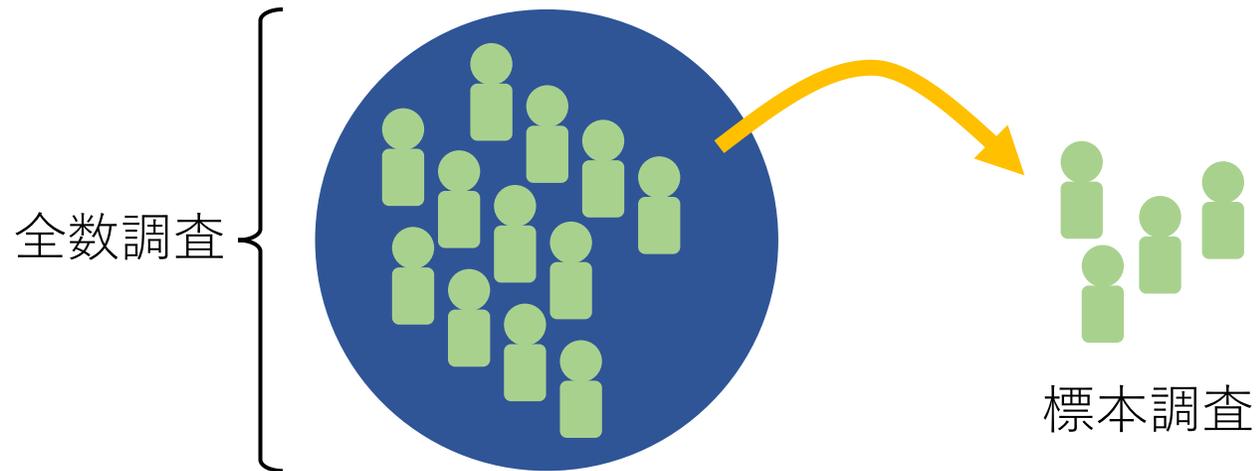
ある集団の特徴を調べるとき、調べる対象となる集団を**母集団**と呼びます。



# 全数調査と標本調査

母集団の全体を調査するのが**全数調査**です。

ただし、母集団があまりにも大きいときには、その全体を調べることは現実的ではありません。このようなとき、母集団の一部を抜き出して調査します。これを**標本調査**と呼びます。



# 標本調査における留意点

母集団の一部を抜き出すことで、母集団が持つ特徴を調べることはできるのでしょうか。標本調査を行う際には、

**無作為抽出**：どのようにして偏りなく標本を抜き出すか

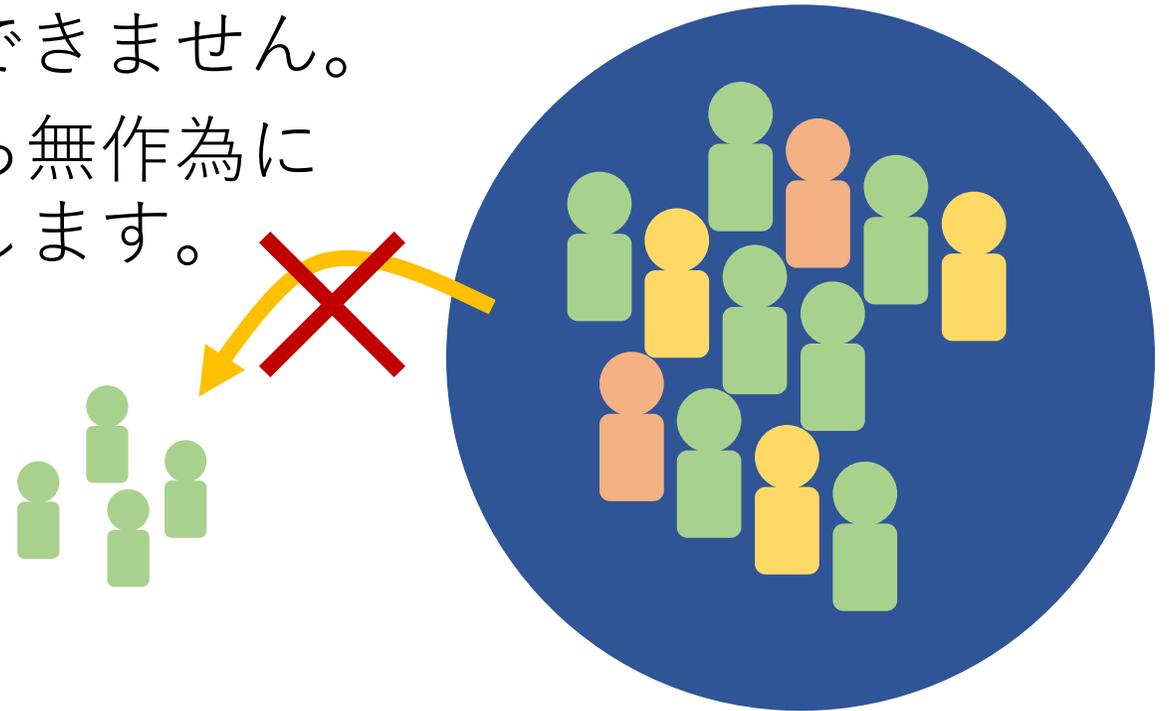
**標本の大きさ**：どれだけの標本を取り出すべきか

に留意しなくてはなりません。

# 無作為抽出

様々な特徴を持つ人から成る母集団で、ある特徴を持つ人に偏った標本調査をしてしまうと、母集団の特徴を推定できません。

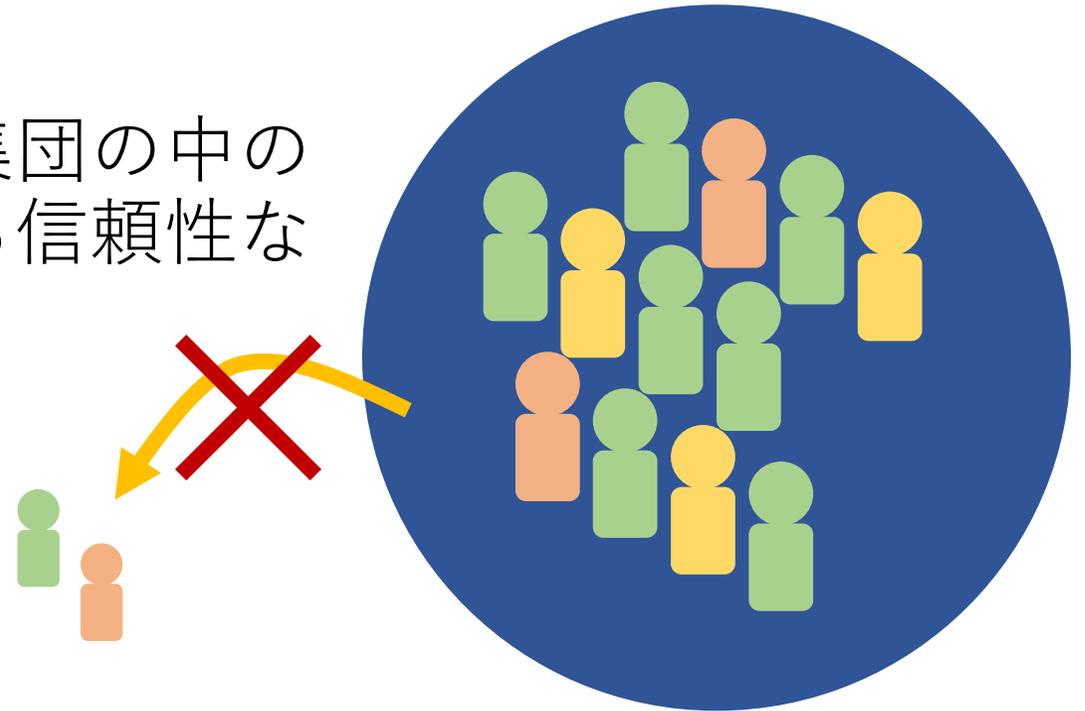
これを避けるには、母集団から無作為に（ランダムに）標本を抜き出します。



# 標本の大きさ（サンプルサイズ）

ランダムに標本を選ぶ場合でも、標本の大きさが小さいと、母集団の特徴を再現できなくなります。

必要となる標本の大きさは、母集団の中のばらつき、得られた結果に求める信頼性などによって決まります。



# 標本の抽出方法

大学生を対象とした調査を例に、無作為抽出について考えます。

**単純無作為抽出**：対象となる大学生に割り振られた学生番号からランダムに数字を選んで標本を抽出します。



**層別抽出**：調査結果に学部毎の差が小さく無いと予想される場合、大学生を学部別（層別）に分けて、学部毎に無作為抽出します。

# 例題

大学生を対象とした学習時間の調査を例に考えます。

20000人の学生から500名を抜き出して標本調査します。また、下表のように、学部毎の学生数はばらつきがあります。

このとき、単純無作為抽出と学部別の層別抽出では、どちらが適切でしょうか。

学部	理	工	農	法	経済	教育	文	医	歯	薬
学生数(人)	3400	4800	1400	1300	1000	4200	1100	1400	800	600

# 解説

単純無作為抽出は、20000人の学生から無作為に500名を抜き出します。例では、工学部の学生数は薬学部の8倍ですが、工学部の学生を薬学部の8倍程度抜き出すとは限りません。

学部別の層別抽出では、学部毎に無作為抽出しますので、学部毎の学生数に合わせて学生を抽出できます。また、学部毎の比較も容易になります。

学部	理	工	農	法	経済	教育	文	医	歯	薬
学生数(人)	3400	4800	1400	1300	1000	4200	1100	1400	800	600

# アンケート調査



アンケート調査では、人々あるいは組織に同じ質問に行い、回答してもらうことで、調査対象に関するデータを集め、そのデータを分析することで有用な情報を得ます。

アンケート調査を行うには、調査の目的に合わせて、調査対象、全数調査か標本調査か、標本調査であれば抽出方法、標本の大きさ、調査項目、調査期間などを検討していきます。

# アンケート調査の質問文

ここで、アンケート調査の質問文について考えます。質問文を作成する際に避けるべきものとして、

- 難解な質問
  - ダブルバーレル質問
  - 誘導質問
- などがあります。



# 難解な質問

回答者に理解しにくい語句が含まれる質問は避けるべきです。

例) NISAについてどう思いますか？

と聞かれても、投資に興味がなくNISAについて知らない回答者には答えようがありません。

# ダブルバーレル質問

一つの質問に2つ以上の事項が含まれる質問は避けるべきです。

例) 昨日、あなたはテレビや動画を見ましたか？

と聞かれた場合、テレビか動画の一方しか見なかった回答者は、はいと答えるべきかいいえと答えるべきか迷うかもしれません。

# 誘導質問

回答者の反応を偏らせてしまうような質問は避けるべきです。

例) あなたはスマートフォンを1日に何時間利用しますか？

例) 大学生のスマートフォン利用時間が問題になっていますが、あなたはスマートフォンを1日に何時間利用しますか？

の2通りの質問では、回答者の反応が変わる可能性があります。

# 問題

妥当な測定が行えるように、質問項目を修正しなさい。

1. NISAについてどう思いますか？
2. 昨日、あなたはテレビや動画を見ましたか？
3. 大学生のスマートフォン利用時間が問題になっていますが、あなたはスマートフォンを1日に何時間利用しますか？