

統計学的標本抽出法

高木 廣文

東邦大学医学部看護学科

国際保健看護学研究室

e-mail: halwin@med.toho-u.ac.jp

http://homepage2.nifty.com/halwin/takagi.html

1

統計学による結果の一般化

例：TVの視聴率調査

- ・日本のすべてのTVに調査用の機器が設置されている訳ではない。
⇒数千万台の内、数百台～数千台程度

Q. ほんの一部の結果から、各番組の視聴率が計算。なぜよいのか？

2

データの取り方の問題

Q. ほんの一部の結果から、各番組の視聴率が計算。なぜよいのか？

何でも集めればよいわけではない！

- 統計学的に正しい方法でデータを集める

標本抽出 **sampling** の問題

3

母集団について

Q. 誰を調べればよいのか？

TVのある全世帯：母集団
概念上－ population
実際上－ universe

4

全数調査と標本調査

- 母集団 universe の全構成員についての調査

- ・全数調査, 悉皆調査

膨大な人的資源, 予算が必要。
国レベルの調査に限られる。

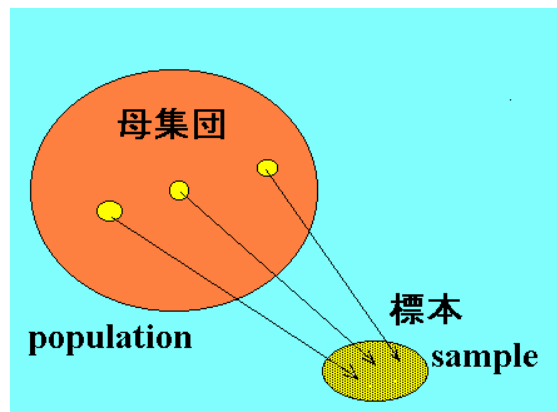
(例) 国勢調査

- 母集団の一部の調査：

⇒ 標本 sample を調べる：「**標本調査**」

5

母集団と標本



標本の選び方は？

Q. どのように標本を選べばよいのか。

- (1) 街頭でやってくれそうな人に依頼する。
- (2) 知人、友人などに依頼する。
- (3) 雑誌などで調査協力者を募集する。
- (4) 会社の職員の家族に頼む。

⇒ すべて「無計画抽出法」×

7

どんな標本が必要か？

Q. 母集団から標本を選ぶのに何が重要か？

母集団を代表するような標本を選ぶ



標本の選び方に偏りが無いこと

8

偏りのない標本の選び方は？

● 等可能性の原則

母集団の各構成員について標本となる可能性が全員等しくなっていること

Q. どのような標本の選び方をすればよいのか？

9

Random に選ぶということ

Q. 正しい標本の選び方とは？

ランダム random に母集団から選ぶ

Q. それではランダムとは何か？

日本語では「無作為」

10

作為と無作為

Q. 「無作為」ではよく分からない。

作為的：故意に手を加えた作り事？



無作為：故意でないこと？

11

無作為に標本を選ぶ

● ランダムな標本抽出：無作為抽出法
⇒ くじ引きの要領で標本を選ぶ。

● 単純無作為（標本）抽出法
simple random sampling
母集団の構成員全員に一連番号をつけ、「乱数 random number」をコンピュータ等で発生させ、番号が一致したのものを標本として選び出す。

12

乱数とは何か？

標本抽出などの目的のために生成される
0～9までの数値から構成される数列で
数値に偏りや特定の傾向が一切存在しないもの

- ・ 0～9までの数値の頻度が等しい。
- ・ 数値のならびに規則性がない。

13

いろいろな無作為抽出法

●無作為標本抽出法

- (1) 単純無作為抽出法
- (2) 系統抽出法
- (3) 多段（階）抽出法
- (4) 層（別）化抽出法

調査のレベル，情報の有無による。

14

単純無作為抽出法

simple random sampling

1. 対象者リストの作成。
2. 一連番号を振る。
3. コンピュータで乱数を発生させる。
番号に一致した者を標本とする。
4. 必要な人数に達するまで繰り返す。

15

系統的抽出法

systematic sampling

1. 対象者リストの作成。
2. 一連番号を振る。
3. 必要な人数で全体の人数を割る。
(抽出間隔の設定)
4. 最初の間隔（1～抽出間隔番）から，
無作為に1人選ぶ（スタート番号）。
5. 抽出間隔ごとに等間隔で標本を選ぶ。

16

多段(階)抽出法

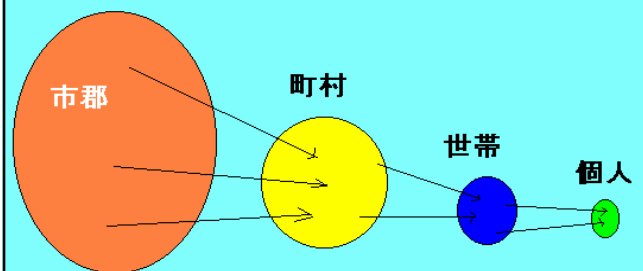
multistage sampling

国レベルの調査のような場合、段階的に
調査対象を絞って行く方法。

1. 人口別に都市を選ぶ。
2. 町村を選ぶ。
3. 世帯を選ぶ
4. 対象者を選ぶ

17

多段抽出法



多段抽出法の考え方(3段階抽出)

層(別)化抽出法

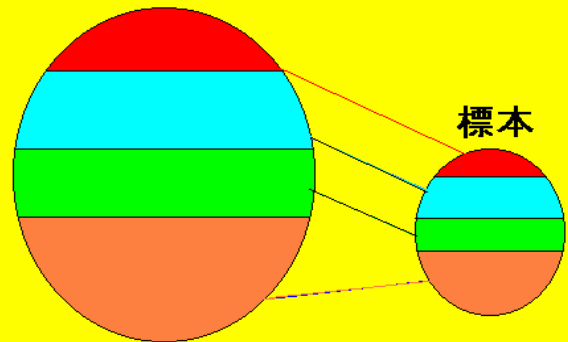
stratified sampling

1. あらかじめ対象集団の特性によって、集団をいくつかの層に分類しておく。
2. 必要な標本数に対して、各層の割合が母集団と一致するように標本を抽出。

最も母集団を代表する標本が得られる。
事前に母集団での特性の分布が必要。

19

母集団



層化抽出法の考え方

臨床研究での標本抽出

- 臨床研究では、完全な無作為抽出は困難。
⇒ 患者データの入手は比較的容易。
例) 心疾患患者の調査研究。
- ・ 患者のみの研究では、患者特有の偏りを評価できない。
⇒ 適当な比較群が必要：
コントロール control

21

コントロールの選び方

- 一般の健常者がよい。病院などでは、適当な外来患者から選ぶことが多い。
 - ・ 比較可能性 comparability を高めるように選ぶ必要がある。
 - ・ 患者と性、年齢、職業などを一致させる。
マッチング matching
 - ・ 「ケース・コントロール研究」
case-control study

22