

第10回 6月26日の講義内容

- § 3-5. サンプリングの方法 (続)
 - 無作為抽出の方法 (続)
 - 多段抽出法
 - 層化抽出法
- § 3-6. 調査の実施
- § 4. 予備的分析

6/26/02 1

多段抽出法 multi-stage sampling

- 母集団から直接ランダムサンプリングするのではなく、母集団を互いに排反する集団に分割して、まず集団をサンプリングする。次に抽出された集団からランダムサンプリングするという抽出法 (二段抽出)。
- 抽出された第一次集団をさらに互いに排反する小集団に分割し、その小集団をランダムサンプリングする、ということが続けていくのが、多段抽出法である。

6/26/02 2

多段抽出法(2)

- (例) 近畿圏から100人抽出
 - 第一段: 大阪、京都、兵庫、滋賀、和歌山、奈良の6府県から一つの府県を抽出
 - 第二段: 第一段で抽出された府県から市町村を抽出
 - 第三段: 第二段で抽出された市町村からランダムサンプリングで100人抽出

6/26/02 3

多段抽出法(3)

- 抽出の原則: 末端が等確率で抽出されること
- 確率比例抽出
 - 第一段階の抽出で、個体数に応じて抽出確率を調整し、第二段階の抽出でランダムサンプリングする方法
- 等確率抽出: 第一段で複数集団の抽出を行う場合
 - 第一段階の抽出では、ランダムサンプリングで抽出。第二段階の抽出で個体数に応じた標本数を抽出。

6/26/02

4

多段抽出法(4)

- 数値例
 - A学部 在籍 300名
 - B学部 " 200名
 - C学部 " 150名
 - D学部 " 400名
 - E学部 " 150名
 - 5学部から2学部を選び、100人抽出

6/26/02

5

多段抽出法(5)

■ 数値例: 確率比例抽出

	抽出確率	第一段	×	第二段	=	最終
- A学部 300名		2/4		50/300		100/1200
- B学部 200名		2/6		50/200		100/1200
- C学部 150名		2/8		50/150		100/1200
- D学部 400名		2/3		50/400		100/1200
- E学部 150名		2/8		50/150		100/1200

6/26/02

6

多段抽出法(6)

■ 数値例: 等確率抽出

	割当数	第一段	第二段	= 最終
- A学部	300名	2/5		
- B学部	200名	$100 \times (200/600)$	2/5	33/200
- C学部	150名	2/5		
- D学部	400名	$100 \times (400/600)$	2/5	67/400
- E学部	150名	2/5		

■ 第一段でどこが抽出されるかによって、最終抽出確率が変化。

6/26/02 7

層化抽出法 stratified sampling

■ 事前と事後を一致させる工夫の一つ。

■ (例) A社社員から50人抽出

- A社社員の学歴構成: 高卒70%、短大卒以上30%
- 母集団(A社社員)からランダム標本を抽出するときの確率分布(事前) 高卒0.7、短大卒以上0.3
- 抽出された標本の学歴構成(事後) 高卒0.5、短大卒以上0.5 ということもありえる。

↓

標本が母集団を代表していないおそれがある。

6/26/02 8

層化抽出(2)

■ 事前と事後の層構成比を一致させるように、標本の割当数を層毎に決め、層毎にランダムサンプリングする。

= 層化抽出法

■ この例では、「学歴」が層(stratum)に対応。

- 高卒層への標本割当数: $50 \times 0.7 = 35$ 人
- 短大卒以上への標本割当数: $50 \times 0.3 = 15$ 人

6/26/02 9

層化抽出(3)

■ 層化二段抽出:

- 層化抽出と多段抽出の組み合わせ
- 世論調査などでよく用いられる

大都市 → 地域ブロック or 都道府県 → 市町村 → ランダムサンプリング

10万人以上の都市 → "

その他 → "

6/26/02 10

§ 3-6. 調査の実施

■ 調査の実施方法

- 自計式: 調査対象者自身で記入
 - 例) 国勢調査
 - Good: (人に知られたくないような) 微妙な問題の調査がやりやすい。
 - Bad: 用語解釈の不一致が生じやすい
- 他計式: 調査員が記入
 - 例) 世論調査
 - Good: 用語解釈の一致・統一が容易
 - Bad: 微妙な問題の調査が難しい(調査員に影響されやすい)

6/26/02 11

§ 3-6. 調査の実施(2)

■ 実施手段

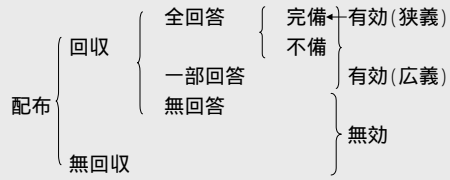
- 人: 訪問
 - 面接
 - 配票
- 郵送 ... 回収率低い(3割程度)
- 電話 ... 対象が特定できない
回答の信頼度が低い
- インターネット ... HTML形式で作成
スクリーニングは容易であるが
対象者は特定できない

6/26/02 12

§ 3-6. 調査の実施(3)

■ エディティング: editing

- 有効票と無効票の区別



6/26/02

13
