

●●● 第9回 12月9日の講義内容

- § 3-5. サンプリングの方法
 - 全数調査と標本調査
 - サンプリングの種類
 - 無作為抽出の方法
 - 単純ランダムサンプリング
 - 系統抽出法
 - 多段抽出法
 - 層化抽出法

1

12/9/09

●●● 全数調査と標本調査

- 統計調査
 - 調査対象 = 母集団
 - **全数調査(悉皆調査)**
母集団全体を調べる(記述)
 - **標本調査**
母集団の部分集合を調べて、母集団の性質をあてる(推測)

2

12/9/09

●●● サンプリングの種類

- 標本のとり方 推定の精度を左右
(標本の代表性)
- 有意抽出(非確率的)
 - 典型調査: 典型的な集団を選ぶ
 - 割当調査: 属性によって調査数を割当
- 無作為抽出(確率的)
 - ランダムサンプリング

3

12/9/09

●●● 無作為抽出の方法

- まず、母集団リスト = サンプリング台帳が必要
 - サンプリング台帳として使用されるもの
 - 「選挙人名簿」
 - 電話帳
 - 各種名簿
 - (注意)個人情報保護法施行以降、サンプリング台帳の利用は制限される方向
 - ➡ 多数のモニターを抱えた調査会社の利用
- サンプリング台帳から確率抽出

4

12/9/09

●●● 単純ランダムサンプリング

- サンプリング台帳から乱数(一様乱数)によって抽出する方法
- 利点: 各個体が等確率で抽出
- 欠点: 母集団の属性構成と抽出後の標本の属性構成が異なることがある。

- 事前確率と事後構成比の違いに注意

5

12/9/09

●●● 単純ランダムサンプリング(2)

12/9講義後追加訂正

- 標本数と精度の関係
 - (例1) 視聴率 関西地区(調査世帯600)
 - プロボクシングWBC世界フライ級タイトルマッチ
内藤大助 × 亀田興毅
毎日放送 '09/11/29(日) 20:15 ~
視聴率 43.1%
 - 二項分布より、視聴率 p の推定量の分散は $p(1-p)/n$
 $p=0.431, n=600$ として計算すると、 0.0004 。
標準偏差は、 0.02 。
したがって、95%信頼区間は $\pm 4.00\%$ 。
- c.f. 視聴率調査について
(視聴率データはビデオリサーチ社の公表にもとづくものである)

6

12/9/09

●●● 単純ランダムサンプリング(3)

- 標本数と精度の関係(続)
(例2) 視聴率 関東地区(調査世帯600)
 - ドラマ 不毛地帯
フジテレビ '09/12/3(木) 22:00 - 54 11.4%
二項分布より、視聴率 p の推定量の分散は $p(1-p)/n$
 $p=0.114$, $n=600$ として計算すると、 0.00016834 。
標準偏差は、 0.0130 。したがって、95%信頼区間は
 $\pm 2.6\%$ 。
 - 視聴率 11.4% と 視聴率 8.8%
の響きの違いは大きい!!

7

12/9/09

●●● 系統抽出法 systematic sampling

- サンプリング台帳から等間隔で抽出する方法。等間隔抽出法ともいう。
- まず、母集団の総数 N を抽出する標本数 n で割り、 $k=N/n$ 個体からなる n 個のサブグループに分ける。
- 次に第1サブグループから乱数によって標本を1つ抽出する。以降サンプリング台帳から k 番おきに抽出する。
- (例) 1,000人の母集団から50人の標本を抽出するケース
 - $k=1000/50=20$ 。
 - サンプリング台帳の1~20番から乱数で r 番目の個体を抽出。
 - 以降、台帳から $r+20 \times j$ 番目 ($j=1, \dots, n-1$)の個体を抽出。

8

12/9/09

●●● 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。

9

12/9/09

●●● 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

10

12/9/09

●●● 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。

11

12/9/09

●●● 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(1200人)から2学部を選び、100人抽出

12

12/9/09

