

2004.4.22 比較現代日本論研究演習 (田中重人)

第2回「統計分析の基礎」目次

1. データ収集から分析まで
2. データの配布
3. 標本抽出

【データ収集から分析まで】

- データの収集 (実験／観察)
- データの特徴を少数の数値に要約して記述 = **記述統計**
- 誤差の評価

(この手続きの一部が**推測統計**)

(教科書 p. 1-6)

【データの配布】

1995 年 SSM 調査 B 票の一部

★ 全国から 70 歳以下の有権者を
層化 2 段無作為抽出

★ 訪問面接法

cf. 『日本の階層システム』（全 6 巻）
東京大学出版会、2000 年。

★ 意識項目と基本的属性に限定

(調査票の×印はデータセットにない項目)

★ 250 ケースをランダムに抽出

★ データが流出しないように

★ 変数ラベルは菅野剛

(日本大学) 氏による

- ★ 毎回の授業で使うので、
忘れないこと
- ★ 期末レポート提出時に返却

【標本抽出の4段階モデル】

ユニバース (universe)

母集団 (population)

計画標本 (designed sample)

有効標本 (valid sample / case)

★ 伝統的な統計学では 4 段階に
わけずに、2 段階で考えるのが
ふつう：

母集団=Universe + population

標本 = (designed/valid) sample

【無作為抽出】

母集団から計画標本を選ぶ際に、
母集団にふくまれるすべての個体
の抽出確率が等しくなるように
抽出する (random sampling)

➡ 「**等確率標本**」

つぎの条件が必要：

★ 母集団の人口が既知

★ 個体を網羅した「台帳」

※ 個体によって抽出確率が違う場合も、事後的に調整して
等確率標本と同様の統計処理をおこなうことは可能

※ 「台帳」が完備してない状況でも、工夫次第で
無作為抽出に近づけることができる

統計的な推測は、**等確率標本を前提とする**

実際の調査で理想的な標本抽出ができることはまずない。

また計画標本のなかから無効回答がでるので、

無作為ではない誤差がかならず発生する。

この誤差は統計的には処理できないので、個別に推測する

- どの層を過剰に代表しているかを把握する
- おなじ母集団を対象にした調査と比較する

【宿題】

論文や新聞・雑誌記事で使われている調査データについて、
標本抽出の4段階にそって紹介する。

人数分コピーを用意してきて、次回授業時に報告。