

第4講 統計分析の基礎 (5/11)

田中重人 (東北大学文学部准教授)

[今回のテーマ] データの性質に関する基本的な事項を理解する

1 データ収集から分析まで

1.1 データの収集

実験 (experiment) とは:

観察 (observation) とは:

1.2 分析可能な形への加工

- 分析の単位
- 変数の同定
- 変数値の付与 (coding)

1.3 データ・セット作成

- データ入力
- クリーニング

2 記述と推測

「統計をとる」ことの第2、第3段階 (教科書 p. 1-6)

- データの特徴を少数の数値に要約 = 記述統計 (descriptive statistics)
- 誤差の評価 (この手続きの一部が推測統計 inferential statistics)

科学的な研究においては、分析結果の「正しさ」についての最終的な決着は、追試の繰り返しによる再現性のチェックによって行われるはずである。しかし、実際には……

- 費用や人的資源の不足などから、頻繁に追試がおこなわれない分野のほうが多い
- 厳密な追試が原理的に不可能であることも多い (歴史的な研究など)

このため、分析結果を公表する際には、誤差に関する情報をできるかぎり公表することが仁義となっている。推測統計は、この目的のために使われる標準的な手法のひとつ。

3 尺度水準

教科書 p. 8

- 比率尺度 (ratio scale)
- 間隔尺度 (interval scale)
- 順序尺度 (ordinal scale)
- 名義尺度 (nominal scale) → 「質的変数」と呼ばれることもある

上位の尺度のほうがあつかえる演算が豊富であり、また下位の尺度の特徴を兼ね備えている

→ 分析手法の選択幅がひろい

私たちが測定するものはたいてい順序尺度以下である (SSM 調査の調査票参照)。

→ 上位の尺度への変換には一定の理論的根拠が必要

実際には、本来は順序尺度のはずの変数について平均値を求めて分析する、といった類のことが広くおこなわれている。