

# 政治学 3（計量政治学）

## 7 回目

秦 正樹

京都府立大学公共政策学部 准教授

hatamasaki@kpu.ac.jp

2022/05/31

火 3 コース@一号館視聴覚室

## 7 回目のテーマ

### ① 7 回目のテーマ

- 「統計的検定」の考え方
  - 2 変量（多変量）の関係と「差分」について考える
  - 記述統計から推測統計に進化させよう！
  - 統計的検定の考え方を徹底的に理解する！
- とにかく「統計的検定」の考え方をしっかりマスターしよう！

### ② 従業準備

- Online Rstudio はここからアクセスしてください。
- Teams にレジユメをあげてありますのでご確認を。
- 資料と online Rstudio を相互に動かしていくので、前回までのコードラインのご準備よろしくお願いします

## 復習：データの種類

### ① 変数と尺度

- 量的変数（比例尺度）：0 を原点として、間隔と比率に意味があるもの  
e.g. 体重：100kg は、50kg に比べて 2 倍重いといえる
- 質的変数（名義尺度）：数値とは無関係に、定義的に区別するためにつけられたもの  
e.g. 性別（1. 男性，2. 女性，3. その他）→数字に性別の意味はない
- 量と質の間：順序尺度：数字の大小には意味があるが間隔が等価でないもの  
e.g. 順位（1 位・2 位・3 位…）．世論調査の多くはこれ？

### ② 変数から大枠を把握する方法

- 量的変数：記述統計を見ながら、大きな傾向を把握  
e.g. 日本人の平均身長，平均年齢（あとでやってみる）と分散など
- 質的変数：平均値に意味はないので，度数分布を見る  
e.g. あるデータの都道府県平均が 17.2 でした→意味不明…

## 宿題: 各政党への感情温度

- 宿題: 各政党への感情温度 (q4\_1\_x) の平均値を確認
  - ✓ 無党派層に最も嫌われている政党はどこでしょうか
  - ✓ 男女で最も好感度に差がある政党はどこでしょうか
  - ✓ 20代とそれ以外の年齢層との間で、最も好感度に差がある政党はどこでしょうか
  - 上の3つを分析して具体的な数値とともに答えを示しておいてください。

## 性別ごとの無党派率を確認

### ● 性別と党派性のクロス表

- 党派性のありなし変数 (q11\_1 を使って pid.dummy とする変数) を作成

→ 政党支持の全体では、政党支持率 \_\_\_\_ % v.s. 無党派率 \_\_\_\_ %

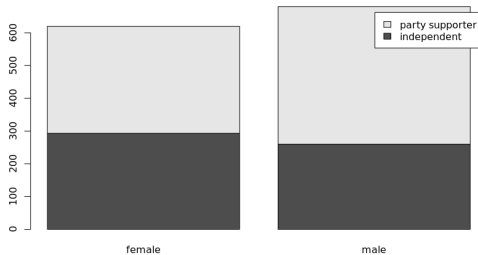
- 党派性のありなし変数 (pid2) を性別とクロス表 table) で示す

### ● 無党派率の性差の関係

- 無党派率：男性 \_\_\_\_ % v.s. 女性 \_\_\_\_ %
- 政党支持率：男性 \_\_\_\_ % v.s. 女性 \_\_\_\_ %
- 無党派は女性のほうが男性よりも \_\_\_\_ %高い？

## この差は偶然？必然？

- 無党派率の性「差」は「意味がある」差なのか？
    - 無党派は女性のほうが男性よりも \_\_\_\_\_ %高い？
    - これらの好感度「差」が、必然的に生じた意味のある差 (=有意差) なのか、それとも意味のない差 (=ただの誤差) なのか問題
- 統計的検定によって、有意差かどうかを検証できる！



## 統計的検定の論理的基盤

### ● 統計的検定の考え方 1：対立仮説と帰無仮説

- 検証したい仮説 (対立仮説) → 「性別と橋下好感度の間には差があるだろう」説  
→ 残念ながら、これを直接に検証する手段がない…そこで逆のパターンを考える!
- 対立仮説と真逆の仮説 (帰無仮説) → 「性別と橋下好感度の間には差がないだろう」  
→ 帰無仮説 (2 つの変数は関係を持たず独立している場合) を基準に考える!

### ● 統計的検定の考え方 2：「敵の敵は味方」理論で考える

- 帰無仮説の方が正しい確率が超高い場合＝対立仮説が間違ってる確率が超高い  
→ 「X と Y には差がない」確率が 99%なら「X と Y に差がある」確率は 1%…
- 帰無仮説の方が正しい確率が超低い場合＝対立仮説が間違ってる確率が超低い  
→ 「X と Y には差がない」確率が 1%なら「X と Y に差がある」確率は 99%!
- \* 敵 (帰無仮説) の方が間違ってる確率が高いやんけ! (ということは、自分の仮説 (対立仮説) が正しい確率が高いってことやろ?) という論理

## 算数でわかる統計的検定の基盤

## ● 実測値と期待度数（理論値）を計算・比較

実際の結果（実測値）

党派性			性別	
	無党派	政党支持者	合計	性別%
	女性	293	327	620 47.7%
	男性	260	420	680 52.3%
	合計	553	747	1300
	党派性%	42.5%	57.5%	

差のない場合（期待度数）

	無党派	政党支持者
女性	263.74	356.26
男性	289.26	390.74

● カイ二乗値を計算： $\chi^2 = \sum (ExpV - ObsV)^2 / ExpV$ 

$$* \chi^2 = (263.74 - 293)^2 / 263.74 + (289.26 - 260)^2 / 289.26 \dots \approx 10.80..$$

	無党派	政党支持者
女性	3.25	2.40
男性	2.96	2.19

● 自由度 (df) を計算： $m \times n$  の分割表における  $df = (m-1) \times (n-1)$ 

$$* \text{ 今回の例で言えば, } 2 \times 2 \text{ なので } df = (2-1) \times (2-1) = 1$$



## カイニ乗分布にもとづく両側検定

- Y 軸が“p-value(p 値)”と呼ばれ、社会科学では慣例的に 5%棄却域が多い
  - カイニ乗分布表 (リンク) から計算可能→  $df=1$  のときの 5%棄却域の  $\chi^2=3.84$
  - 近年、アメリカ統計協会の声明があって、p 値の基準は今後変化する可能性大
  - この分析では、 $\chi^2=10.80$  なので、Y 軸:p=0.001014 となる
- 帰無仮説が正しい確率はほぼ 0.1%であるので、私の方が 99.9%の部がある

