

2002.7.5 比較現代日本論研究演習 I (田中重人)

## 第10回「平均値と標準偏差」

1. 代表値と散布度
2. 平均値
3. 標準偏差
4. SPSS のコマンド
5. 平均値を使うときの注意事項

# 【代表値と散布度】

平均値 (mean) - 標準偏差 (SD)  
(間隔尺度以上)

中央値 (median) - 四分位偏差 (Q)  
(順序尺度以上)

(教科書 p. 42–51)

# 【平均値】

総和をデータ数で割ったもの

# 【標準偏差】

平均値からの偏差の2乗値の平均が「分散」  
分散の平方根が「標準偏差」

平均値と標準偏差はセットで使う

# 【SPSS のコマンド】

「記述統計」 「記述統計」

変数とオプションを指定

# 【平均値を使うときの注意事項】

平均値ははずれ値の影響を受けやすい。

あまりにかけはなれたケースがあるときは

- ・ 上下数%を取りのぞいたデータセットで計算する (調整平均：教科書 p. 46)
- ・ 順位に変換したり中央値を使って分析

平均値・標準偏差は**間隔尺度**以上のデータ  
に対してしか意味をもたない。

順序尺度の平均値をとっていいのは

- ・ 潜在的には間隔尺度のはず
- ・ 測定のポイントが一定間隔

という2条件をともに満たす場合

(4/26 資料参照)

- ➡ 具体的には  
4点以上の尺度  
正規分布に近似 (教科書 p. 53–59) :
- 単峰性
  - 左右対称性 (歪度)
  - 中央への集中度 (尖度)

ヒストグラムを描いて検討するとよい。

正規分布との乖離度を統計的に検討する手法もある

- ➡ これらの条件を満たさない場合は  
非線形変換 (教科書 p.142–144)  
順位に変換したり中央値を使って分析



間隔尺度のデータでも、  
**左右対称でないもの**については  
平均値よりも中央値のほうが  
適当であることが多い

典型例：収入・人口など

次のデータの平均と SD は?

{0, 1, 4, 5, 7}

## 【課題】

適当な変数について、ヒストグラムの軸上に  
平均値と標準偏差を（手書きで）書き入れた  
ものを作って提出