

現代日本学演習 II 「統計分析の基礎」

第7講 平均と分散

田中重人 (東北大学文学部准教授)

[テーマ] 平均値と標準偏差の定義と計算

1 代表値と散布度

教科書 pp. 42-52 を読んで、「中央値」「四分位偏差」「平均」「標準偏差」の計算方法を理解する。
特に、表 2-1 (p. 48) で何が計算されているかを考えること。

重要なこと：

- 代表値と散布度はセットで使う
- 尺度水準による制限

2 平均値と標準偏差

平均 (mean): 総和をデータ数で割ったもの

分散 (variance): 平均値からの偏差の2乗値の平均

標準偏差 (standard deviation): 分散の平方根 (SD と書くことが多い)

教科書の表 2-1 (p. 48) を例にとると、いずれも3人分のデータセットが3つあり、

A: 2, 3, 4

B: 1, 3, 5

C: 3, 3, 3

これらのデータのそれぞれについて、つぎの計算をする

- 平均値を求める
- 「偏差」(p. 49) を求める
- 2乗する (これはなぜか)
- 全部足す (「平方和」(sum of squares: SS) という)
- 合計人数 (=3) で割る (これが「分散」)
- 平方根をとる (これが「標準偏差」)

3 宿題

教科書 p. 52 の練習問題 2-3 について、平均値と標準偏差を計算せよ。計算の途中経過がわかるように解答すること。