

第9講 分散分析 (つづき)

田中重人 (東北大学文学部教授)

[テーマ] 分散分析 (ANOVA) の考えかたと計算方法を理解する

1 前回課題について

男性: 1, 2, 3, 3, 4 → 平均 2.6

女性: 2, 3, 4, 4, 5 → 平均 3.6

全体の平均 (SD): 3.1 (1.14)

グループ別平均値を当てはめた「仮想」データの平方和は、つぎのようになる。下線部に注意。

$$\text{仮想の平方和} = 5(\underline{2.6 - 3.1})^2 + 5(\underline{3.6 - 3.1})^2 = 2.5 \quad (1)$$

これを $N (=10)$ で割って平方根をとると標準偏差が得られる。

$$\text{仮想SD} = \sqrt{\frac{2.5}{10}} = 0.5 \quad (2)$$

$$\eta = \frac{\text{仮想SD}}{\text{実際のSD}} = \frac{0.5}{1.14} = 0.44 \quad (3)$$

ただし、SPSS では平方和を $N - 1 (=9)$ で割って「標準偏差」を求めているので、注意 (前々回資料参照)。相関比 η を求める場合は、分子・分母の両方がおなじ方式で計算できていれば問題ない。

仮想SDのことを「群間」のSDと呼ぶことがある。一方、各グループ内での平均からの偏差を使って求めた標準偏差を「群内」のSDと呼ぶ。

$$\text{実際のSD}^2 = \text{仮想SD}^2 + \text{群内のSD}^2 \quad (4)$$

分散分析の実際の計算では、平方和どうして割り算して η を求める (N で割らずに済み、平方根を求めるのも一度で済むため)。結果を示す「分散分析表」は、伝統的にはそのようにして作成する。

教科書 pp. 203-208 の説明では SS_A が群間の平方和に、 SS_E が群内の平方和に、それぞれ相当する。

2 表の書きかた

- 各層と全体の平均値と標準偏差 (測定水準の2桁下まで)
- 各層と全体の人数
- 相関比またはエフェクトサイズ (小数第3位まで)
- 欠損数とその原因

3 グラフの書きかた

平均値をプロットし、上下にSDを表示する。誤差範囲 (error bar; 別名「ヒゲ」) にはSD以外を書く場合もあるので、必ず「±標準偏差」であることを明記する。

Excel では

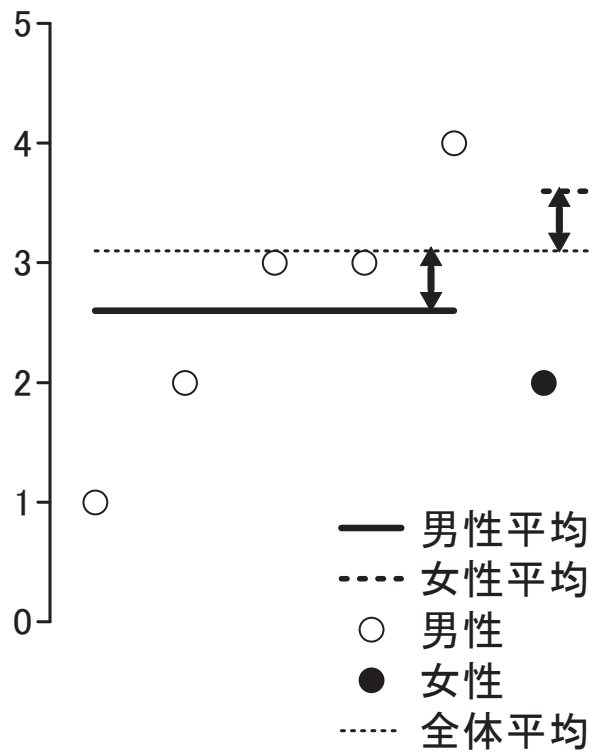
- PSPP 出力を (HTML 形式でファイルに書き出して) シートにはりつける
- 折れ線グラフを描く
- メニューの「レイアウト」から「誤差範囲」→「その他の誤差範囲オプション」をえらぶ
- 「ユーザ設定」→「値の指定」
- 「正の誤差の値」「負の誤差の値」にSDが入っているセル範囲を指定 (おなじものでよい)

Google スプレッドシートでは同様のことはできない模様。

より詳細に分布の違いを検討したいときは、グループ別に折れ線グラフを描いてもよい。

文献

奥村 晴彦 (2016) 「効果量, Cohen's d, 検出力, 検出限界」 <<https://oku.edu.mie-u.ac.jp/~okumura/stat/effectsize.html>>



各グループ内の偏差

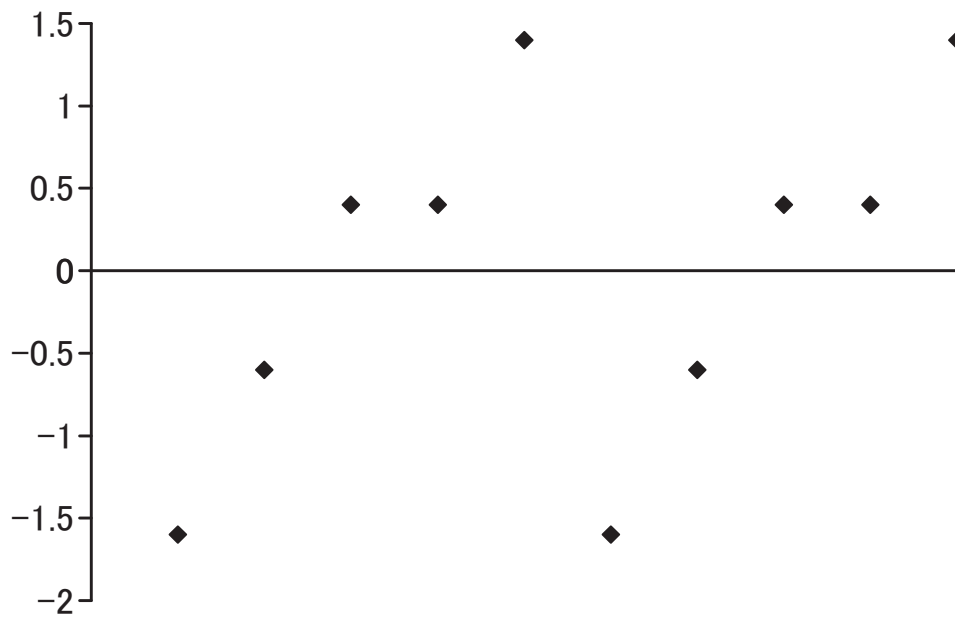
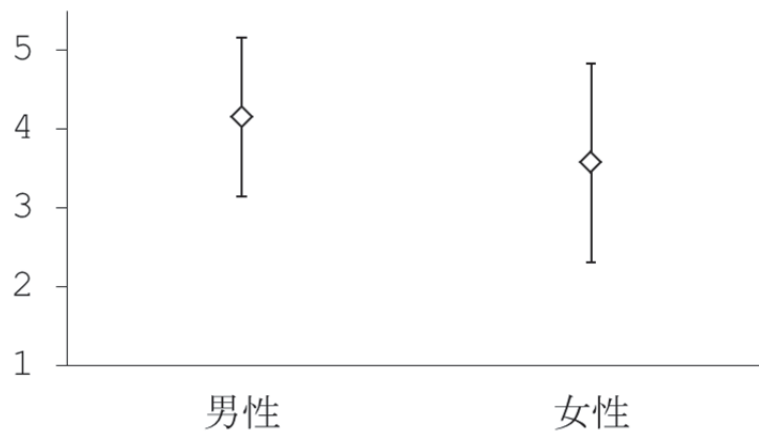


表 1 保守的意識の男女差

	平均	標準偏差	(人)
男性	4.15	1.01	(109)
女性	3.57	1.26	(130)
合計	3.83	1.18	(239)

「以前からなされていたやり方を守ることが、最上の結果を生む」
に対する回答: 「1. そう思う」～「5. そう思わない」
相関比 $\eta=0.244$. 無回答=11.



「以前からなされていたやり方を守ることが、最上の結果を生む」
に対する回答: 「1. そう思う」～「5. そう思わない」
相関比 $\eta=0.244$. N=239. 無回答=11.

図 1 保守的意識の男女差 (平均±標準偏差)