

第9講 分散分析

田中重人 (東北大学文学部准教授)

[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

- 選択肢が4つ以上の項目を選ぶこと
- 正規分布に近似しているか確認すること (単峰性、偏り、集中度)
- グラフの縦軸の単位
- (ESの計算で) 割り算を間違えない

2 分散分析の考えかた

グループ別の平均値を当てはめて仮定の分散を求める分析法を「分散分析」(ANOVA: ANalysis Of VAriance) という。

- 従属変数 (dependent variable) と独立変数 (independent variable)

相関比 (イータ) の性質:

- 最小値:
- 最大値:

大きさの評価基準は、Cramerの連関係数 V と同様。

なぜ相関比を求めると、平均値を比較していることになるのか?

3 課題

次のデータ (10人) について、分散分析を行なう

男性: 1, 2, 3, 3, 4

女性: 2, 3, 4, 4, 5

まず手計算 (またはExcel) で考えてみて、そのあと、SPSSにデータを入力して検算する。

- (1) 全体の平均値とSDを求める
- (2) 男女別の平均値を求める
- (3) 男性の平均値×5人と女性の平均値×5人からなる仮想データを考えてSDを求める
- (4) (3)のSDを(1)のSDで割ったものが相関比 η

この相関比がなぜ「平均値の比較」の指標になるかを考えること。

4 相関比とエフェクトサイズの関係

相関比 η とエフェクトサイズ ES の間にはつぎの関係がある (n_1, n_2 は各グループの度数、 $N = n_1 + n_2$ は全体の度数)。

$$ES^2 = \frac{\eta^2}{1 - \eta^2} \times \frac{N^2}{n_1 n_2} \quad (1)$$

特に、2 グループの度数が等しい ($n_1 = n_2$) なら、この式は次のようになる。

$$ES^2 = \frac{4\eta^2}{1 - \eta^2} \quad (2)$$

(グループの度数が違えば、ES はこれより大きくなる)

さらに、 η があまり大きくない ($\eta < 0.4$ 程度) 場合であれば、次のような単純な式で近似できる：

$$ES = 2 \eta$$

5 モデルとデータの乖離

相関比 η は、モデルとデータの乖離を表した値と解釈できる

- 「モデル」は何か？
- データとの乖離はどうやって計算しているか？
- 係数の取りうる値の範囲は？

6 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

7 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

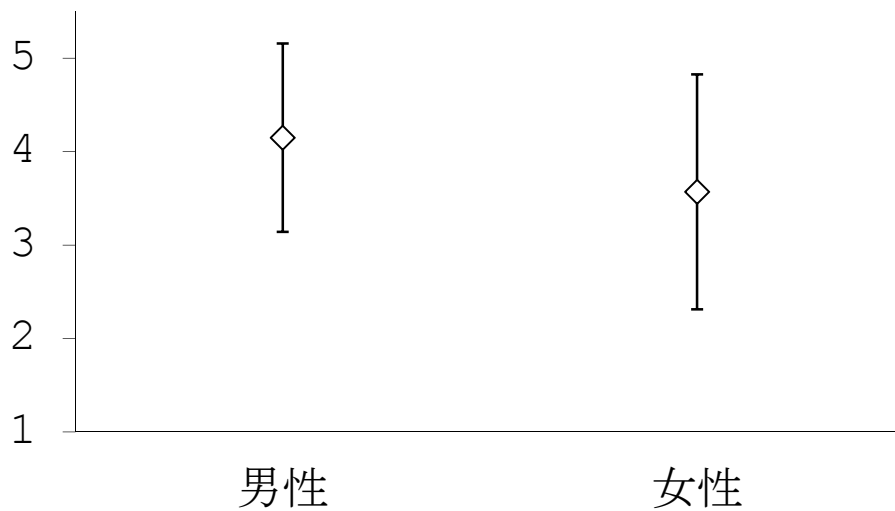
- SPSS 出力をシートにはりつける
- 折れ線グラフを描く
- メニューの「レイアウト」から「誤差範囲」→「その他の誤差範囲オプション」をえらぶ
- 「ユーザ設定」→「値の指定」
- 「正の誤差の値」「負の誤差の値」にSDが入っているセル範囲を指定 (おなじものでよい)

より詳細に分布の違いを検討したいときは、グループ別に度数ポリゴン (または折れ線グラフ) を描いてもよい。

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)