Rによる
反復測定分散分析

PD　井関龍太

心理データ解析演習 (2009/5/13)

反復測定データの分析上の問題

- **反復測定データ**・・・1人の参加者が複数の条件を経験している（**被験者内**要因）
  - 各水準のデータ間に**相関**が生じることが多い
    ⇒F分布に歪みを生じる（タイプⅠエラー変動）
    分散分析の結果が適切でなくなる
- **分散分析**における配慮
  - 主分析・・・球面性の検定と自由度の調整
  - 単純主効果の検定・・・誤差項の選択
  - 多重比較・・・手法の選択，誤差項の選択

球面性の仮定

- F値が正確なF分布にしたがうための**必要十分条件**
  - cf. 分散の等質性の仮定・・・十分条件
- 対応のあるすべての水準対のデータの差の分散が等しい ※水準数が2のときには常に成り立つ
  - **球面性の仮定**：C’ΣC＝λI
    C＝直行正規対比行列，Σ＝母共分散行列，
    λ＝スカラー数，I＝単位行列
- 変量の分布を集中楕円面で表したとき，すべての方向に分布が等しい（球形）状態

Mauchlyの球面性検定

- MauchlyのW統計量
  - 0～1の値をとる
    0＝球面性が成り立たない
    1＝球面性が成り立つ
  - 極めて複雑な分布

$$w=\frac{\left|C'\Sigma C\right|^{N/2}}{\left[\left(1/p\right)trC'\Sigma C\right]^{N_p/2}}$$

※pは反復測定要因の水準数

- χ²分布に近似させて有意性を判定

$$\chi^2=-(N-1)\cdot d\cdot\ln(W)$$

$$df=\frac{p(p+1)}{2}-1$$

$$d=1-\frac{2p^2+p+2}{6(N-1)p}$$

εによる自由度の調整

- 球面性検定が有意 or 球面性の仮定が満たされていない恐れがある
  - F分布に歪み
    ⇒歪みを調整して適切な検定を行う
- **ε（イプシロン）統計量**を使って，分散分析の自由度を調整して検定する
  - たとえば，F (2, 28) = 4.25 なら，
    ⇒F (ε ×2, ε ×28) = 4.25 として，
    調整後の自由度で有意性を判定する

2つのε統計量

- Greenhouse-Geisserのε
  - 0～1の値をとる
    0＝×球面性
    1＝○球面性
  - ε≧.75のときには厳しすぎるとの指摘

$$\hat{\varepsilon} = \frac{\left(\sum_{i=1}^{r} s_{ii}\right)^2}{r\sum_{i,j=1}^{r} s^2_{ij}}$$

- Huynh-Feldtのε
  - 1を超えることがある
    (その場合，1と見なす)
  - 緩めの調整 (Greenhouse-Geisserのεの値が大きいときにはより適切)

$$\tilde{\varepsilon} = \frac{(Nr\hat{\varepsilon}-2)}{r(N-g-r\hat{\varepsilon})}$$

単純主効果の検定における誤差項の選択

- **プールした誤差項**：すべての水準のデータを使って推定した誤差項
  - 自由度大＝検出力が高い
  - “なるべく多くのサンプルから推定する”という原則に忠実
- **水準別誤差項**：データを分割し，関連する部分のみから推定した誤差項
  - 容易に自由度を調整できる
  - 異質な（可能性のある）分散を合成しない

※球面性に配慮する立場からは，水準別誤差項を推奨（Keselman & Keselman, 1993）

多重比較法の選択

- 大部分の多重比較法（Tukey法，Ryan法とその改良版など）・・・各水準のデータが**独立**であることを前提とする
- 反復測定データ・・・各水準間に**相関**あり
⇒<u>ほとんどの多重比較法が適用できない</u>
- 球面性が成り立たない状況でのタイプⅠエラー率（Keselman & Keselman, 1988; Maxwell, 1980）
  - WSD法＝4.5～11.5%／4.1%～9.5%
  - Bonferroni法＝1.8～6.4%／3.0～5.0%

Bonferroniの方法

- 最もオーソドックスな多重比較法のひとつ
  - 有意水準を**比較の数**で割る
  - **Bonferroniの不等式**：α／k
    3水準の比較（5％）：0.05／3＝0.0167
    4水準の比較（5％）：0.05／6＝0.0083
  - 各仮説を**調整後の有意水準**で評価
- **長所**：球面性のくずれに比較的頑健（データに相関があってもBonferroniの不等式は成り立つ；永田・吉田, 1997）
- **短所**：検出力が低い

Holmの方法

- Bonferroniの方法の**改良版**
  - p値の低い順に仮説を評価する
  - kの値を**ステップごとに1ずつ減らす**
    4水準の比較（5%）：
    - ステップ1：0.05／6＝0.0083
    - ステップ2：0.05／5＝0.0100・・・
  - ただし，順番に評価し，有意でなくなった時点で終了（その後のステップは差なしとする）
- **長所**：検出力が高い（HSD法と同程度）
- **短所**：特になし

ステップダウン法

Shafferの方法

- Holmの方法の**さらなる改良版**
  - 仮説の**論理的関係性**に基づく調整
    (早見表かプログラムを使う)
    4水準の比較(5%)のステップ2:
    - 同時に成立しうる残りの仮説は5つではない
      「$\mu_1 = \mu_2 = \mu_3 \neq \mu_4$」だとすると,
      残りは$\mu_1 = \mu_2$, $\mu_1 = \mu_3$, $\mu_2 = \mu_3$の**3つ**
- **長所**:検出力が高い(Holmより確実に上)
- **短所**:論理的関係性が仮定できないデータには使えない

Shafferの方法のための早見表

- 永田・吉田 (1997, p. 97) の表6.11より

	比較する群の数			
ステップ	3	4	5	6
1	3	6	10	15
2	1	3	6	10
3	1	3	6	10
4		3	6	10
5		2	6	10
6		1	4	10
7			4	7
8			3	7
9			2	7
10			1	6
11				4
12				4
13				3
14				2
15				1

多重比較における誤差項の選択

- 通常，多重比較では，**すべての水準のデータ**から推定した誤差項を用いる（分散分析の主効果・単純主効果の誤差項）
  - この方法は**球面性の仮定**に依存する（Keselman & Keselman, 1993）
- **水準別誤差項**の推奨：各ステップで**比較する2水準ごと**の誤差項を用いる（ペアワイズ比較）

多標本球面性の仮定

- **混合要因計画**：反復測定要因に加えて，独立測定要因（**被験者間**要因）を含む
  - 独立測度の水準（群）ごとに反復測定要因の相関関係が異なる可能性
    ⇒**群の等質性**も考慮しなければならない
- **多標本球面性の仮定**：混合要因計画における（反復測定要因を含む効果の）F値が正確なF分布にしたがうための条件
  - 球面性＋共分散行列の等質性の仮定

Mendozaの多標本球面性検定

- 多標本球面性の仮定を検定するには
  - **二段階**：共分散行列の等質性の検定をしてからMauchlyの球面性検定を行う
  - **一段階**：2つの検定を一度に行う
- Mendoza (1980) による検定方法：一段階多標本球面性検定
  - MauchlyのWの代わりに，**λ***を計算する
  - λ*は0～1の値をとる
  - χ^2分布に近似させて検定を行う

Rによる実行：ANOVA君の利用

- R・・・オープンソースの統計パッケージ
  - 多様な推測統計・多変量解析に対応
  - 分散分析のための関数も複数ある
    ⇒一挙に単純主効果の検定，多重比較まで行いたい（既存の関数以上の要求）
- ANOVA君・・・分散分析用に作成した関数
  - 要因計画法に沿った使用を想定
  - 球面性に配慮した検定方式
  - デフォルトでは，Mendozaの検定（Mauchlyの検定もオプションで指定可能）

ホームページ

ANOVA君の使用法（1）：ソースコードの読み込み

1. Rを起動する
2. Rのメニューバーから「ファイル」をクリック
3. 現れたメニューの中から「Rコードのソースを読み込み」をクリック
4. ファイルの選択画面になるので，ANOVA君のファイルを選んで「開く」をクリック（このとき，ファイルの種類を「All files」にすること）

→Rのコンソールに「source(“C:～」といったコードが表示される（特にエラーメッセージなどが出なければ，読み込み成功）

ANOVA君の使用法（2）：データファイルの作成

- データは，以下の形式で並べる
  - 2×2×3の混合要因計画（AsBC）の例：

被験者内要因

被験者間要因

	b1			b2		
	c1	c2	c3	c1	c2	c3
a1	4	7	5	4	3	2
a1	6	3	6	3	4	3
a1	8	6	4	5	4	4
a2	4	8	9	6	7	2
a2	2	2	8	1	3	2

個々のデータ

被験者間要因の水準（群）を表すラベル列

ANOVA君の使用法（3）：データの読み込み

1. 読み込むデータをマウスでドラッグし，右クリックして「コピー」を選択
2. Rのコンソールに以下のコードを入力して「Enter」

```
> dat <- read.table(“clipboard”)
```

※“dat”の部分は任意の変数名

ANOVA君の使用法（4）：関数の実行

- コンソールにコマンドを入力して「Enter」

```
> anovakun(データ, "要因計画の型", 各要因の水準数, ...)
```

  - 「データ」・・・データを読み込むときにつけた変数名。
  - 「要因計画」・・・分析したい計画の型。「A」～「Z」を各要因を表すラベル，「s」を被験者のラベルとする。被験者間要因は「s」の左側，被験者内要因は「s」の右側に配置する。
  - 「各要因の水準数」・・・「A」～「Z」の各要因の水準数を順に入力する。

参考文献

- 千野直仁 (1993). 反復測度デザイン概説－その1－　愛知学院大学文学部紀要, 23, 223-236.
- 千野直仁 (1994). 反復測度デザイン概説－その2－球形検定とその周辺についての批判的レビュー－　愛知学院大学文学部紀要, 24, 103-119.
- 千野直仁 (1998). 反復測定 (測度) 分散分析／基礎と応用 http://www.aichi-gakuin.ac.jp/~chino/
- Huynh, H. (1978). Some approximate tests for repeated measurement designs. *Psychometrika*, 43, 161-175.
- Keselman, H. J., & Keselman, J. C. (1988). Repeated measures multiple comparison procedures: Effects of violating multisample sphericity in unbalanced designs. *Journal of Educational Statistics*, 13, 215-226.

- Keselman, H. J., & Keselman, J. C. (1993). Analysis of repeated measurements. In L. K. Edwards (Ed.), *Applied analysis of variance in behavioral science* (pp. 105-145). New York, Marcel Dekker, Inc.
- Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. *Journal of Educational Statistics*, 5, 269-281.
- Mendoza, J. L. (1980). A significance test for multisample sphericity. *Psychometrika*, 45, 495-498.
- 永田靖･吉田道弘 (1997). 統計的多重比較法の基礎　サイエンティスト社
- 入戸野宏 (2004). 心理生理学データの分散分析　生理心理学と精神生理学, 22, 275-290.