統計学(基礎)

第9回
一元配置分散分析

分散分析(ANOVA)

- 分散分析(ANOVA:ANalysis Of VAriance)
  - 「全体のばらつき」を「群間分散」と「郡内分散」の2つに分けて考える
  - 群間分散(Between-group)
    - グループの平均値どうしの差がどれくらいあるか→ グループ分けによる違い
  - 群内分散(Within-group)
    - 同じグループ内でデータがどのくらいバラついているか→ 個人差や誤差
  - F値 = 群間分散 ÷ 群内分散

分散分析（ANOVA）

- 平均の違いを「分散の比」として扱う検定
  - 複数の平均の違いをまとめて判断
  - 3群以上の場合に（t検定の拡張）
- F値 ＝ 群間分散 ÷ 群内分散
  - 群間分散が大きい：グループ平均が離れている
    → 群による違いが大きい
  - 群内分散が大きい：同じグループ内の個人差が大きい
    → グループ内がバラバラ
    - 群間分散が大きく、群内分散が小さいほど、グループ間の違いが有意

t検定と分散分析の関係

- t検定は平均値の差を見る検定
  - 「平均の差 ÷ 標準誤差」でt値を計算
  - A群とB群の平均がどれくらい違うかを直接比較
- 2群の比較の場合：$F(1, n_1 + n_2 - 2) = t(n_1 + n_2 - 2)^2$
  - t検定と2群の一元配置分散分析は同じ検定統計量を別の形で見ているだけ。
  - 検定結果(p値)は完全に一致。
  - といって、2群はt検定でやることになっている

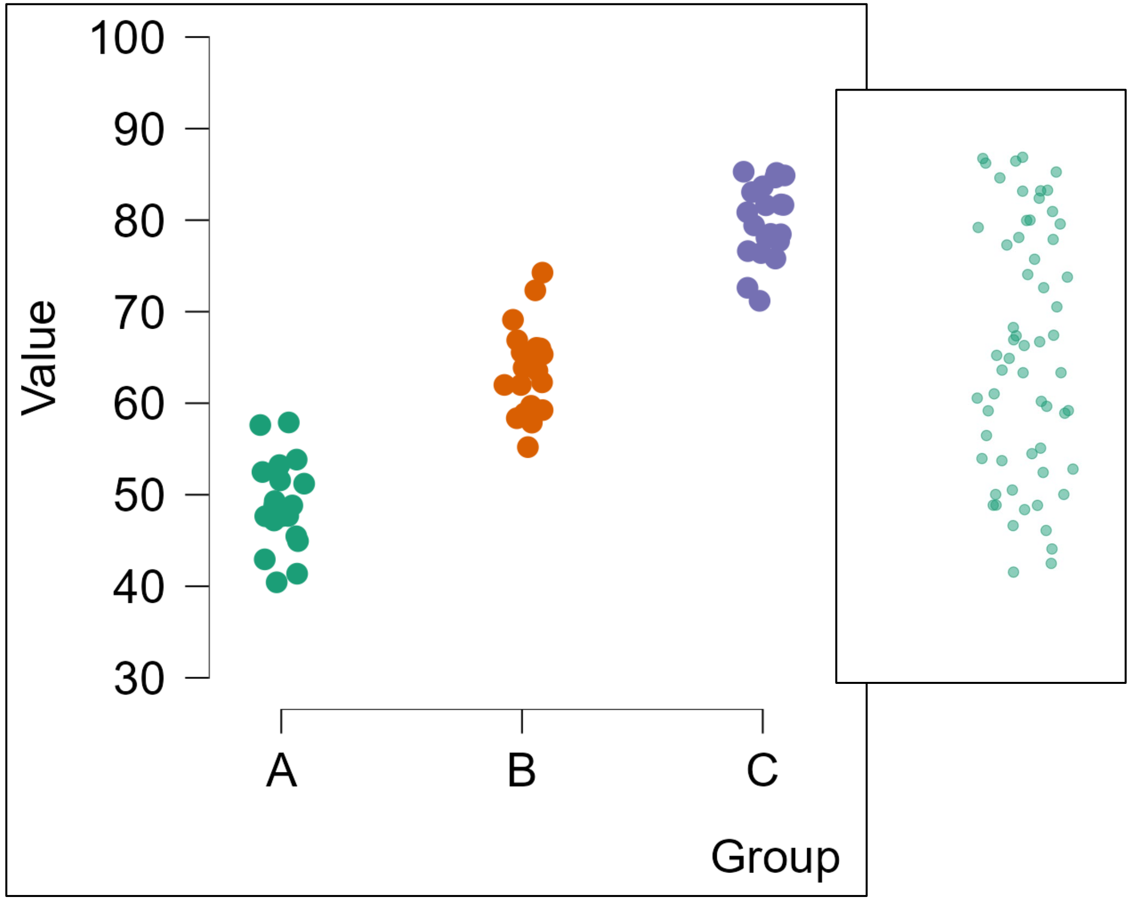
なぜt検定を繰り返してはいけないか

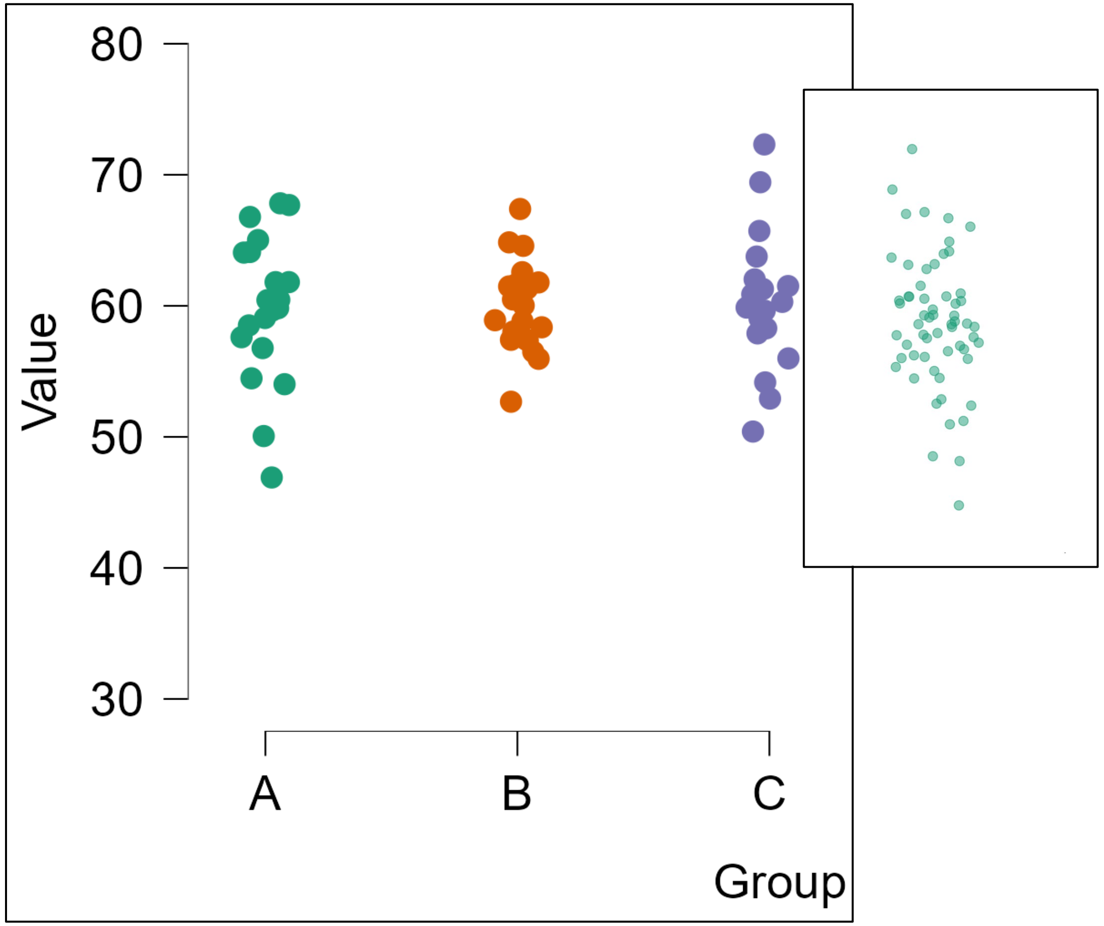
- 第一種の過誤(Type I Error)の問題
    - 7回目(t検定)参照
  - 3群だと、3回のt検定
  - 4群だと6回のt検定
  - n群の場合　n(n-1)/2　　${}_{n}C_{2}$

分散分析の仕組み

- 例:3つのクラスのテスト
  - クラスごとの平均点が大きく違う
    → クラス間のばらつき(群間分散)が大きい
  - 同じクラスの中で点数がバラバラ
    → クラス内のばらつき(群内分散)が大きい

分散分析の仕組み


100
90
80
70
60
50
40
30
Value
A
B
C
Group


80
70
60
50
40
30
Value
A
B
C
Group

分散分析の仕組み

- 分散分析（ANOVA）は、「全体のばらつき」を次の2つに分けて考える
  - 群間分散（Between-group）
    - 群の平均値どうしの差がどれくらいあるか
      → グループ分けによる違い、群間の差
  - 群内分散（Within-group）
    - 同じ群内でデータがどのくらいバラついているか
      → グループ内の個人差や誤差

分散分析の仕組み

- 群間分散 ÷ 群内分散 ＝ F値
  - 群間が大きく、群内が小さい
    → Fが大きい → グループの平均が違うっぽい
  - 群間も群内も同じくらい
    → Fが小さい → 違いはなさそう
  - 群間の違いが大きく、群内のばらつきが小さいほどF値が大きくなる
- 分散分析は、「全体のばらつきのうち、どれだけがグループ分けによって説明できるか」を調べる方法。

群間分散・群内分散・F値

- 群間分散(MSB)
  - SSB:群間平方和(Sum of Squares Between)
  - MSB:群間平均平方(Mean Square Between)
  - 分母の k−1は自由度(kは群の数)

$$SS_B = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$

$$MS_B = \frac{SS_B}{k-1}$$

- 群内分散(MSW)
  - SSW:群内平方和(Sum of Squares Within)
  - N:全体のデータ数

$$SS_W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$$

$$MS_W = \frac{SS_W}{N-k}$$

- F値

$$F = \frac{MS_B}{MS_W}$$

分散分析を行うための条件

- 独立性
  - 各群のデータは独立している
  - 実験計画で確保するしかない
- 正規性
  - 各群の母集団が正規分布
  - nが十分大きければ中央極限定理で影響が小さい
- 等分散性
  - 各群の分散が等しい
  - 群サイズが等しければ多少の違いは問題になりにくい

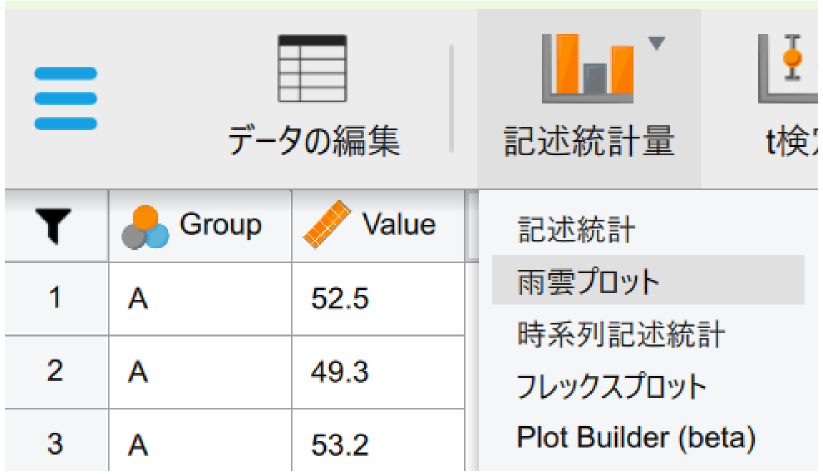
分散分析を行うための条件

- 分散分析は、「多少の正規性の崩れ」や「分散の不均一」にはロバスト
- 特に各群のサンプルサイズが同じくらいなら、かなり安定
  - 群の大きさが大きく違う、外れ値が多いときは影響を受けやすい

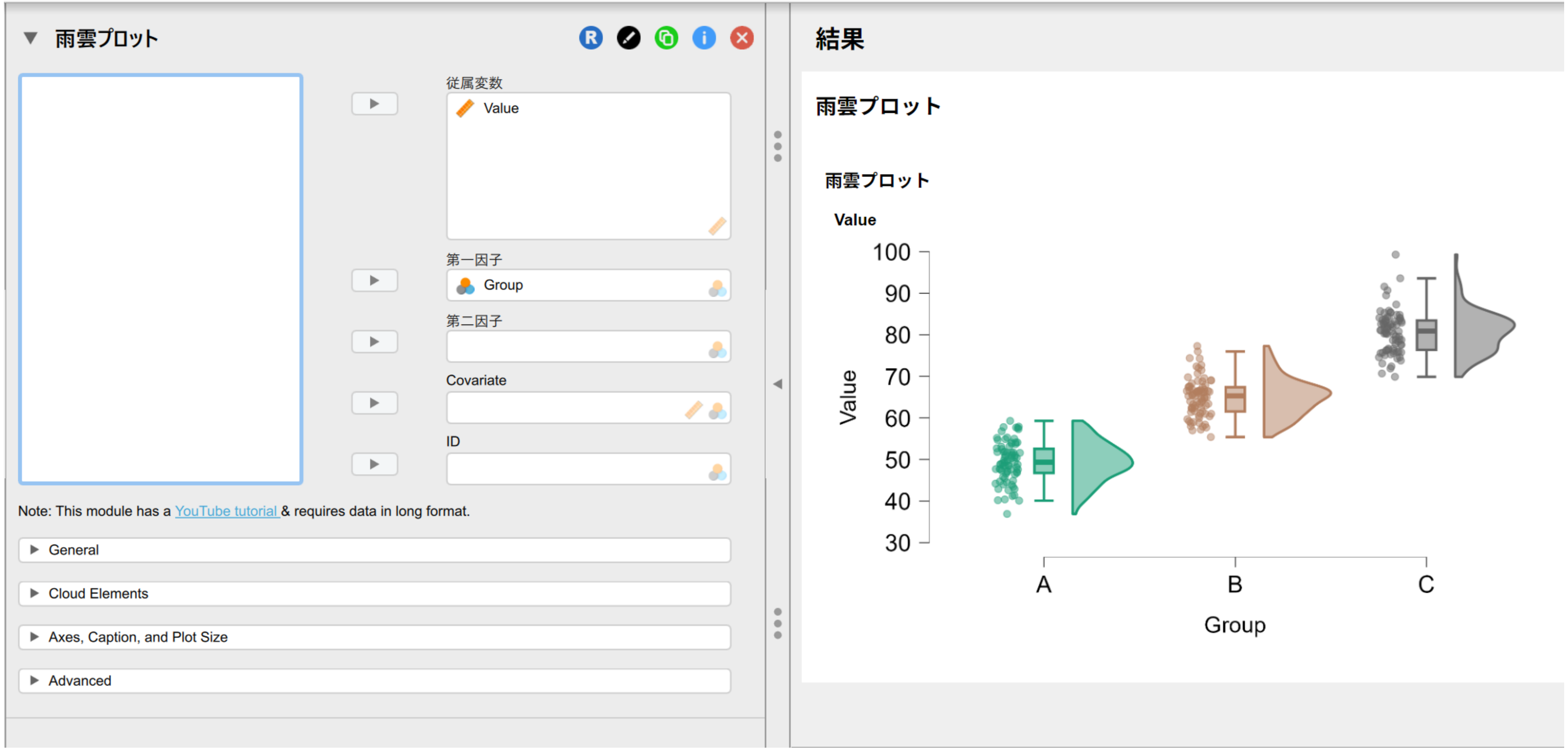
- 正規性や等分散性の検定はあるけど、今はしないでプロットで考える
  - t検定と同じ

雨雲プロット(Raincloud plots)

- 2019年に提唱
  - Raincloud plots: a multi-platform tool for robust data visualization(T. Marshall, M. Allen, D. Poggiali, R.A. Kievit, 2019)
    - https://pmc.ncbi.nlm.nih.gov/articles/PMC6480976/?utm_source=chatgpt.com
- JASPは採用に積極的



雨雲プロット(Raincloud plots)



雨雲プロット（Raincloud plots）

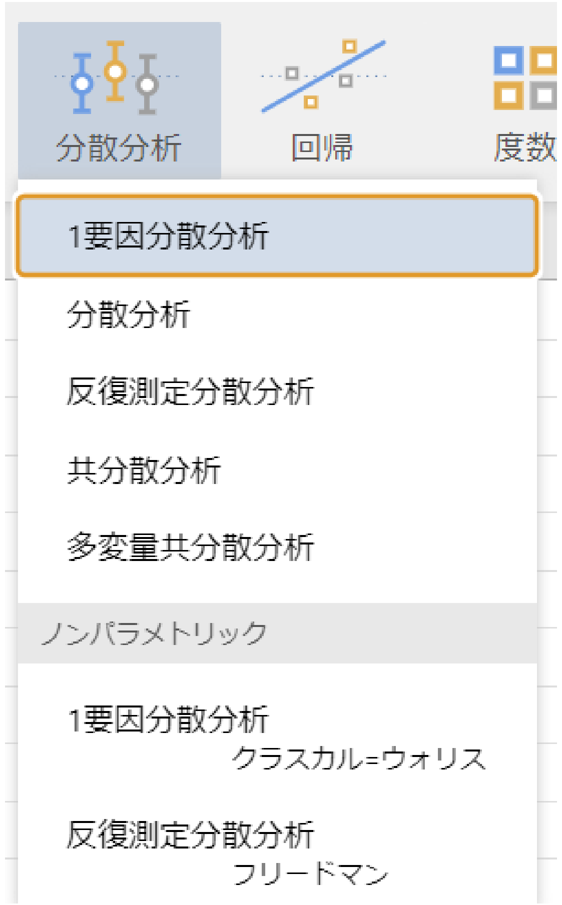
- 今のところJASPは標準で搭載
- JASPは統合的な環境に積極的（SPSSっぽい）
  - 出力もAPA準拠
- jamoviはシンプルな環境（保守的・初学者向け）を目指す

分散分析実行の実際

- 左 JASP 右 jamovi

※jamoviには、
「1要因分散分析」と
「分散分析」がある
どっちも結果は一緒

- data09_01



分散分析 JASP

分散分析

分散分析 - Value

ケース	平方和	df	平均平方	F	p
Group	39085	2	19542.32	831.3	< .001
Residuals	5572	237	23.51		

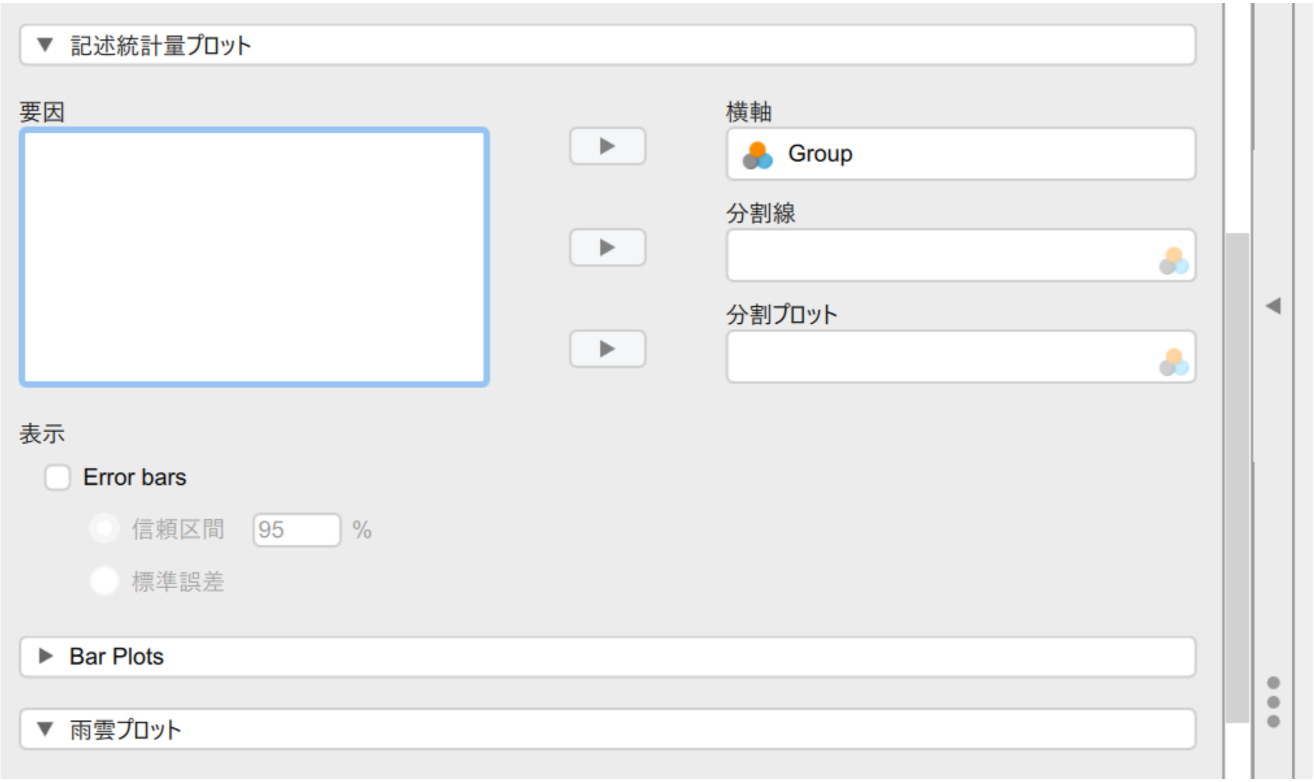
注 タイプ III 平方和

記述統計量

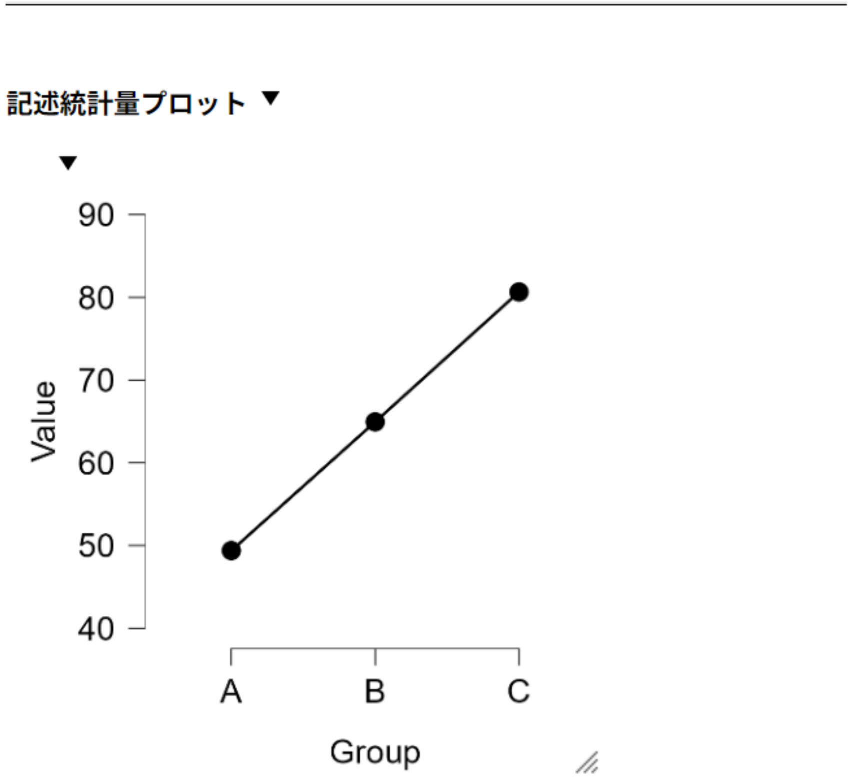
記述統計量 - Value

Group	N	平均値	標準偏差	標準誤差	変動係数
A	80	49.38	4.788	0.535	0.097
B	80	64.94	4.587	0.513	0.071
C	80	80.64	5.154	0.576	0.064

分散分析 JASP

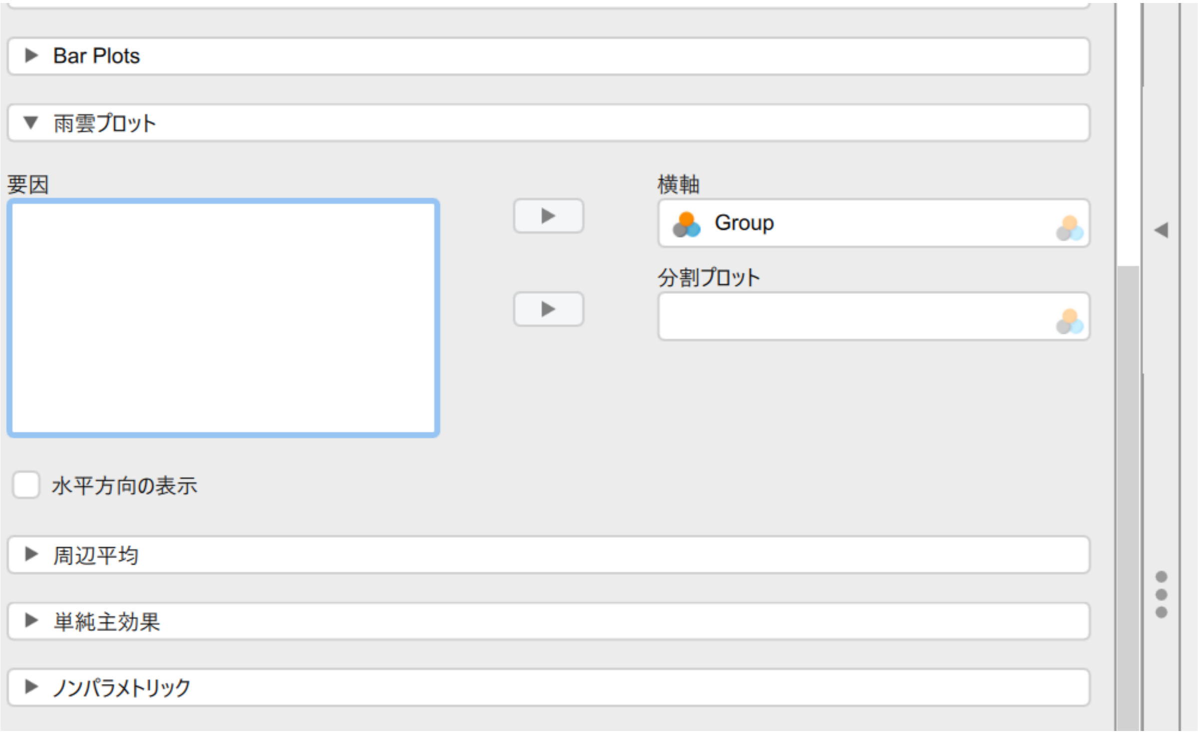


記述統計量プロット
要因
横軸
Group
分割線
分割プロット
表示
Error bars
信頼区間
95
%
標準誤差
Bar Plots
雨雲プロット

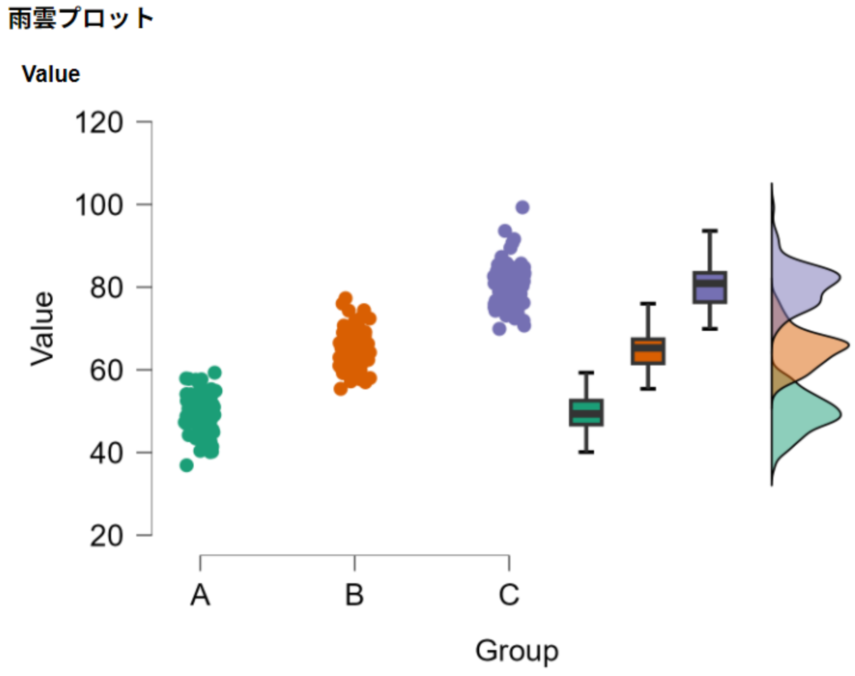


記述統計量プロット
Value
90
80
70
60
50
40
A
B
C
Group

分散分析 JASP



Bar Plots
雨雲プロット
要因
横軸
Group
分割プロット
水平方向の表示
周辺平均
単純主効果
ノンパラメトリック



雨雲プロット
Value
120
100
80
60
40
20
Value
A
B
C
Group

1要因分散分析　jamovi

1要因分散分析

1要因分散分析（フィッシャー法）

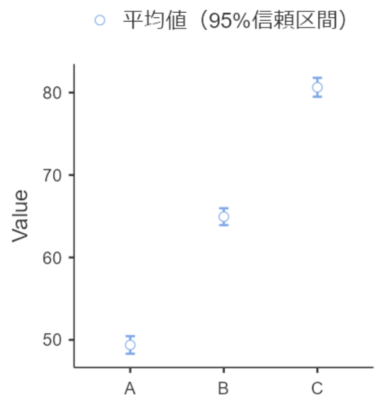
	F	自由度1	自由度2	p
Value	831	2	237	< .001

グループ統計量

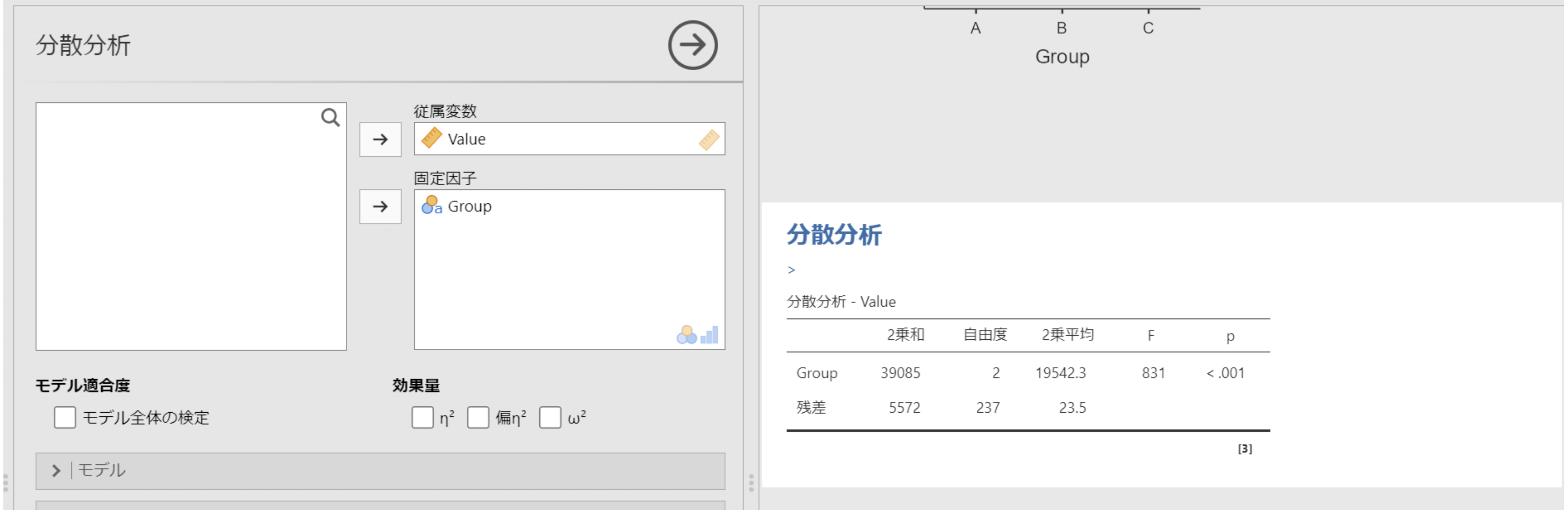
	Group	N	平均値	標準偏差	標準誤差
Value	A	80	49.4	4.79	0.535
	B	80	64.9	4.59	0.513
	C	80	80.6	5.15	0.576

グラフ

Value

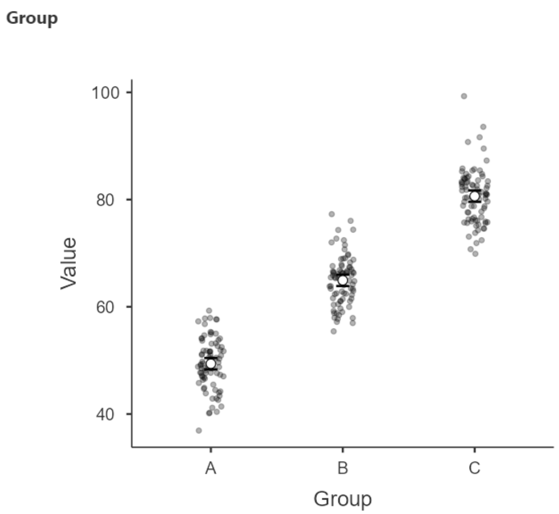


分散分析　jamovi


分散分析
従属変数
Value
固定因子
Group
モデル適合度
モデル全体の検定
効果量
η²
偏η²
ω²
モデル
A
B
C
Group
分散分析
分散分析 - Value
2乗和
自由度
2乗平均
F
p
Group
39085
2
19542.3
831
< .001
残差
5572
237
23.5
[3]

分散分析　jamovi

推定周辺平均



推定周辺平均値 - Group

グループ	平均値	標準誤差	95%信頼区間 下限	95%信頼区間 上限
A	49.4	0.542	48.3	50.5
B	64.9	0.542	63.9	66.0
C	80.6	0.542	79.6	81.7

分散分析と多重比較

- 群の平均値の違い（群間分散）が、各群内のバラつき（群内分散）に比べてどれだけ大きいかをF値で見る
- 群にしたことによる差はわかっても、どこに差があるかはわからない
→多重比較（事後の多重比較）

多重比較

- 全体で群に分けたときに差があるかどうか見る
  →分散分析
- 各群間に差があるかどうか見る
  →多重比較
  - 分散分析と多重比較は独立した別の検定（結果独立）
    - 理論的には分散分析で有意差があっても多重比較で有意差が無いという状況はあり得る（その逆もある）
    - ANOVA は「平均がどこか1つでも違うか」を見る「全体の検定」
      多重比較は「ペアごとに差を見る」→解像度が違うため、結果が一致しないことがある

多重比較

- まずは分散分析を行う
  - 群で分けた効果があるか(差があるか)を確認
- 差があったら多重比較を行う
  - どこに差があるかを見る

- いきなり多重比較だけをするのではない

多重比較

- 基本的にはTukeyをやればよい
  - 独立群・等分散が前提
    - ロバストなので、多少等分散が怪しくても大丈夫
    - 群サイズが大きく異なる場合は Games-Howell の方が望ましい
  - 歴史的にANOVA向けになっているので使いやすい
    - 分散分析で差があるのに、Tukeyで差が出ないことになりにくい
    - Tukey は、複数の平均を比べたときに「いちばん大きく見える差」の確率に合わせて、有意の基準を決める。その分だけハードルが上がるため、どのペア比較でも全体の誤検出率(Type Ⅰ Error)が5%以内に収まる。比較全体の最大値を使って安全側に調整している。(Bonferroniは有意水準を群数で割り算して調整)

事後の多重比較　JASP

B	80	64.94	4.587	0.513	0.071
C	80	80.64	5.154	0.576	0.064

記述統計量プロット

雨雲プロット

事後検定

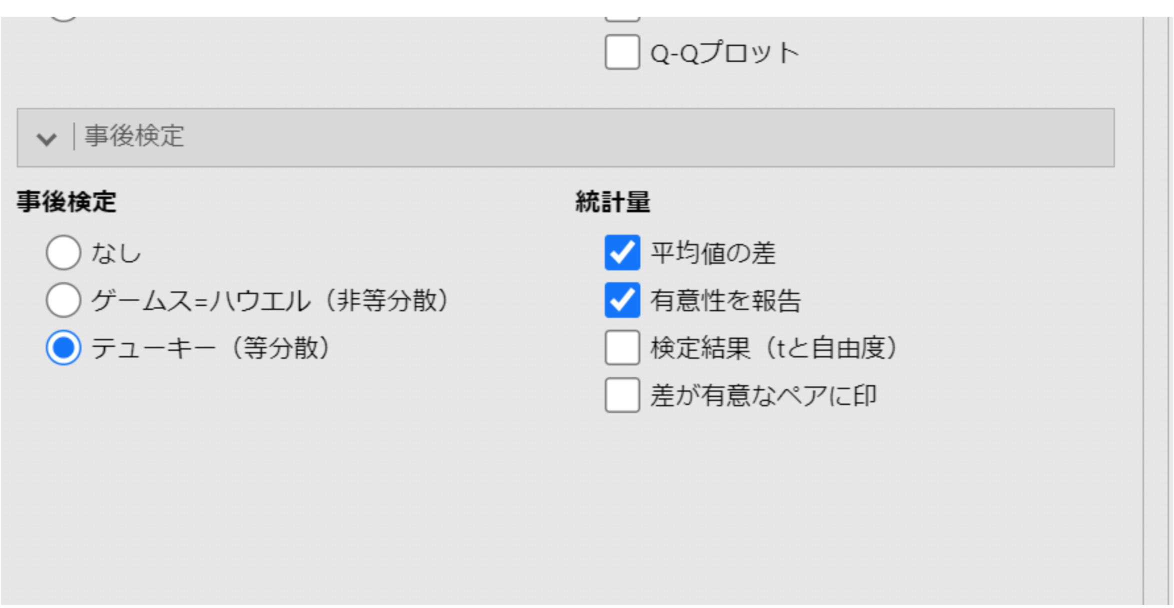
標準 (HSD)

事後比較- Group

		平均値差	標準誤差	df	t	$p_{テューキー}$
A	B	−15.56	0.767	237	−20.30	< .001
	C	−31.26	0.767	237	−40.77	< .001
B	C	−15.70	0.767	237	−20.48	< .001

注 3 の推定値の族を比較する調整された p 値

事後の多重比較 jamovi(1要因分散分析)

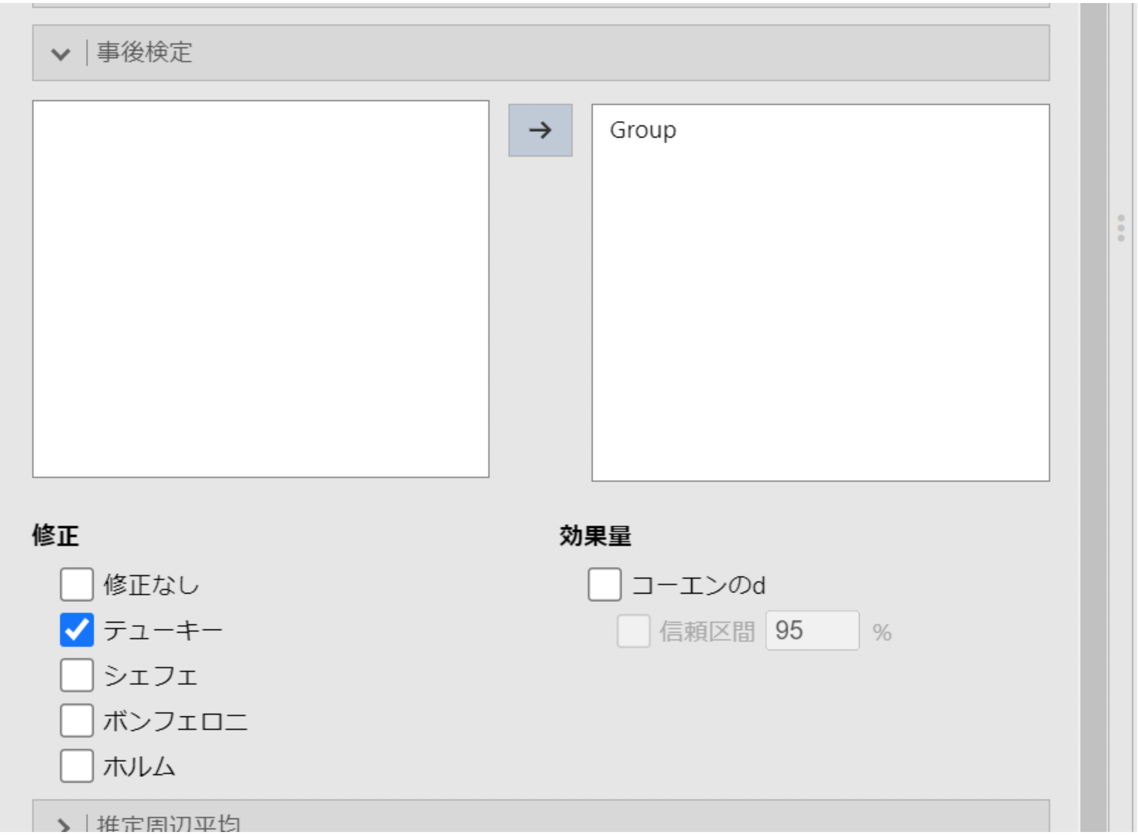


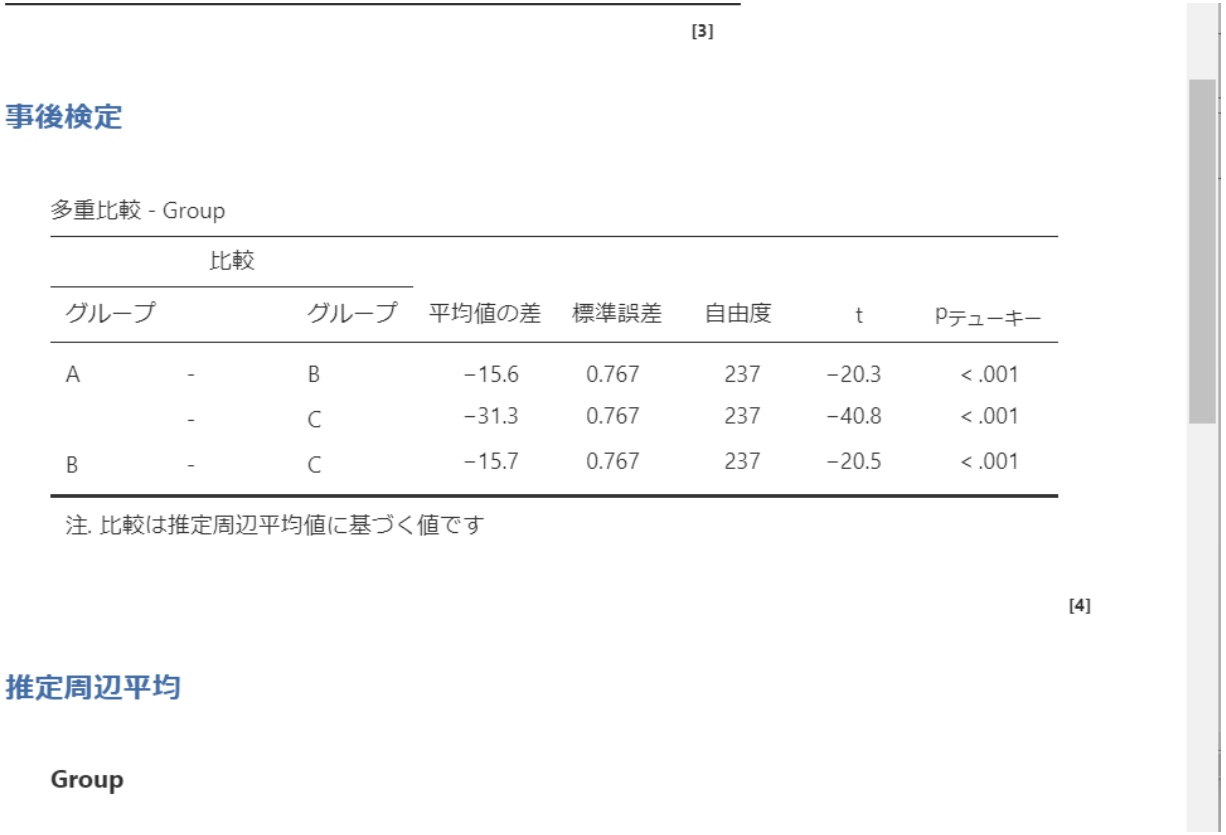
事後検定

多重比較(テューキー法) - Value

		A	B	C
A	平均値の差	—	−15.6	−31.3
	p値	—	< .001	< .001
B	平均値の差		—	−15.7
	p値		—	< .001
C	平均値の差			—
	p値			—

事後の多重比較　jamovi(分散分析)





[3]

事後検定

多重比較 - Group

比較							
グループ		グループ	平均値の差	標準誤差	自由度	t	pテューキー
A	-	B	−15.6	0.767	237	−20.3	< .001
	-	C	−31.3	0.767	237	−40.8	< .001
B	-	C	−15.7	0.767	237	−20.5	< .001

注. 比較は推定周辺平均値に基づく値です

[4]

推定周辺平均

Group

対応のある分散分析ですね

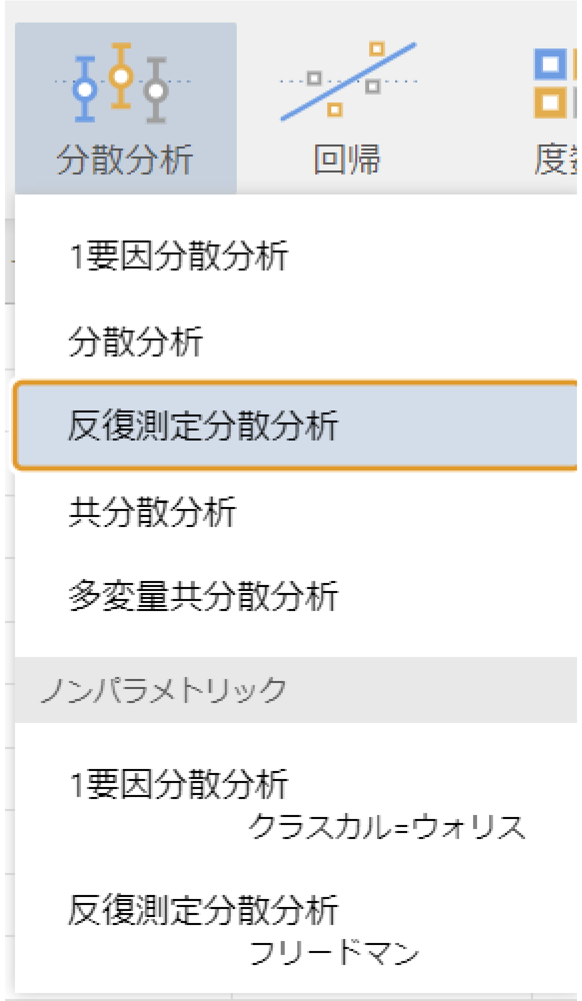
反復測定の分散分析

反復測定の分散分析

- 同じケースで、複数回測定などしている場合
  - 対応のあるt検定の拡張
  - 基本的な考え方は、対応がない場合と同じ
  - データはWide形式

反復測定の分散分析

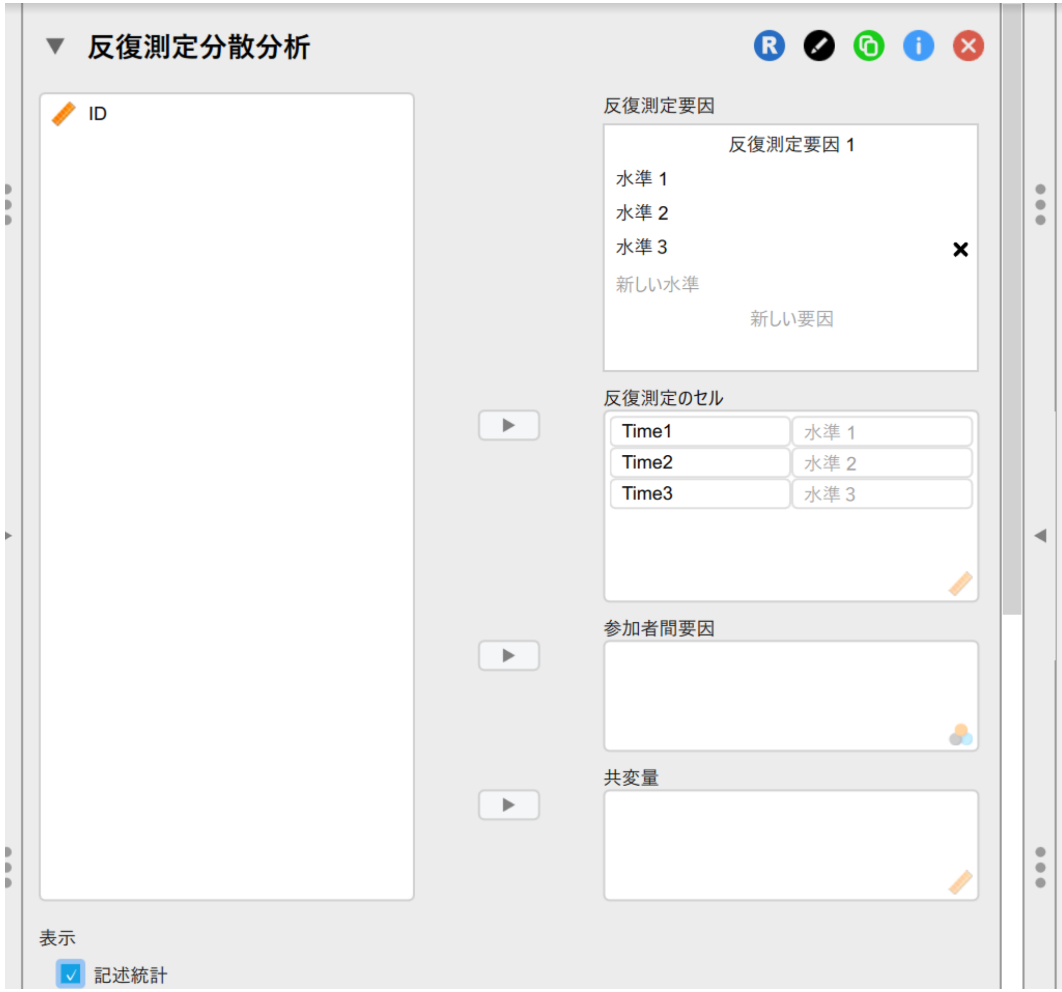
- 左　JASP　右　jamovi
- data09_02



注意点

- どちらのソフトも「反復測定要因(因子)」が2水準なので追加が必要(データがWide形式なので指定が必要)

反復測定の分散分析　JASP



反復測定分散分析

参加者内効果

ケース	平方和	df	平均平方	F	p
反復測定要因 1	18078	2	9038.82	349.1	< .001
Residuals	4091	158	25.89		

注 タイプ III 平方和

参加者間効果

ケース	平方和	df	平均平方	F	p
Residuals	1481	79	18.75		

注 タイプ III 平方和

記述統計量

記述統計量

反復測定要因 1	N	平均値	標準偏差	標準誤差	変動係数
水準 1	80	49.38	4.788	0.535	0.097
水準 2	80	59.94	4.587	0.513	0.077
水準 3	80	70.64	5.154	0.576	0.073

球面性の補正 JASP

反復測定分散分析 ▼

参加者内効果

ケース	球面性の補正	平方和	df	平均平方	F	p
反復測定要因 1	None	18078	2.000	9038.82	349.1	< .001
	グリーンハウス・ゲイザー(Greenhouse-Geisser)	18078	1.961	9220.26	349.1	< .001
Residuals	なし	4091	158.000	25.89		
	グリーンハウス・ゲイザー(Greenhouse-Geisser)	4091	154.891	26.41		

注 タイプ III 平方和

参加者間効果 ▼

ケース	平方和	df	平均平方	F	p
Residuals	1481	79	18.75		

注 タイプ III 平方和

記述統計量

記述統計量

反復測定要因 1	N	平均値	標準偏差	標準誤差	変動係数
水準 1	80	49.38	4.788	0.535	0.097

球面性とは

- 球面性（Sphericity スフェリシティ）
  - どの2時点の『変化量のばらつき』も同じである状態

Var(A−B) = Var(A−C) = Var(B−C)

【球面性がある場合】
A→B：ばらつき 中
A→C：ばらつき 中
B→C：ばらつき 中

【球面性が崩れている場合】
A→B：ばらつき 大
A→C：ばらつき 中
B→C：ばらつき 中

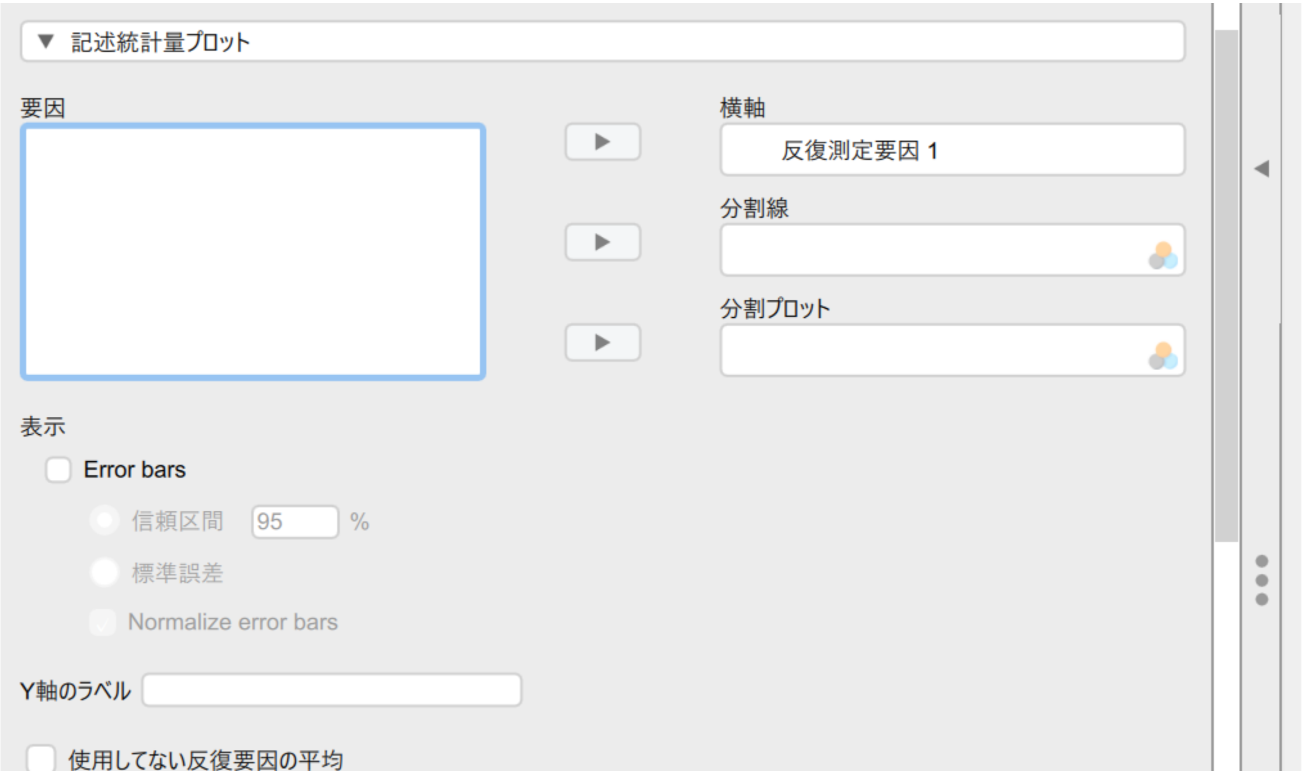
なぜ球面性が必要なのか

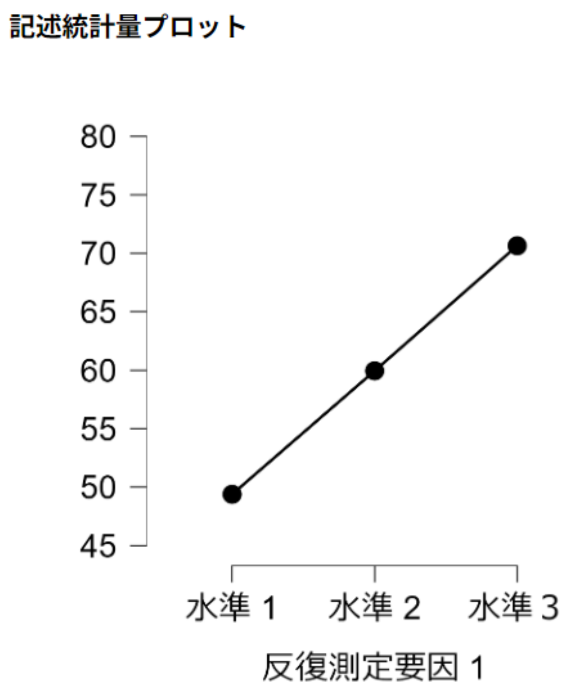
- 球面性が崩れる
  → F 値が過大になりやすい
  → 誤って「差がある」と判断する危険(偽陽性)
- Greenhouse–Geisser(GG)がこの偏りを補正

最初からGreenhouse–Geisser を見る

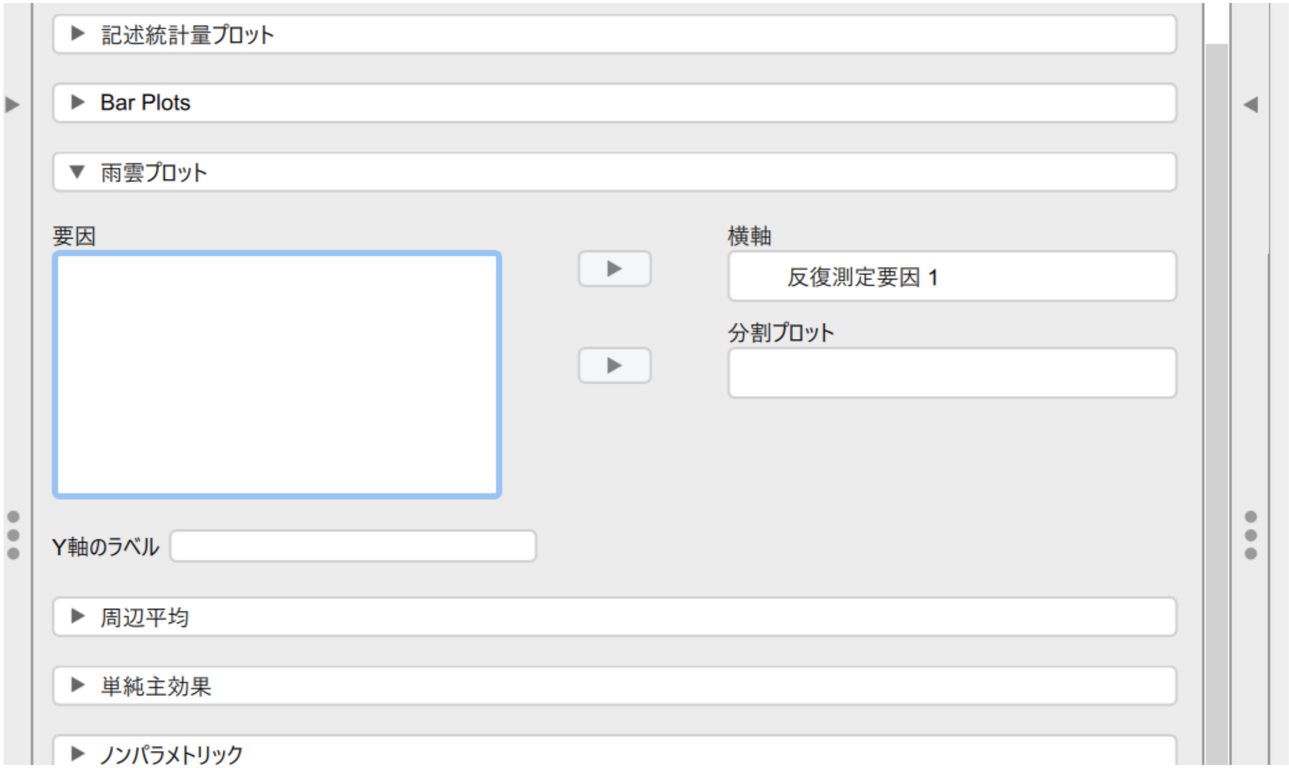
- 本来はMauchly検定を行うことになってはいるが
  - Type Ⅰ Errorを避けたい
  - 球面性が成り立っている場合 GG と None（なし） の差はほぼない
  - 球面性が崩れた場合はGG補正
    - Mauchly検定 の p が < .05 → 球面性なし
- 論文投稿等で球面性検定が必須の場合のみ、Mauchly検定（球面性検定）を実施
  - 通常は検定せずにGGの補正を指定

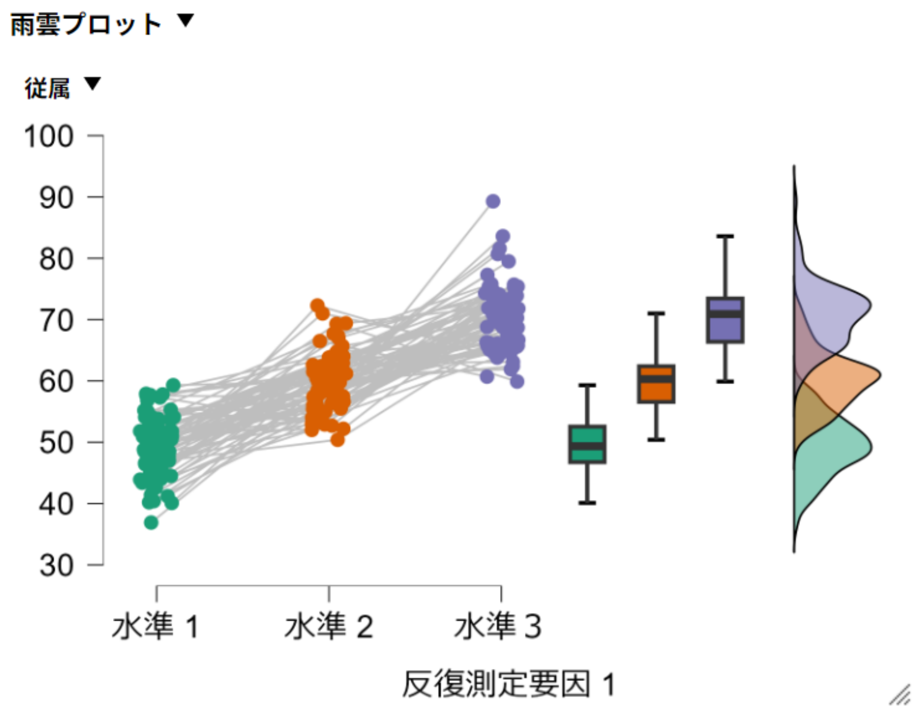
反復測定の分散分析 JASP





反復測定の分散分析 JASP



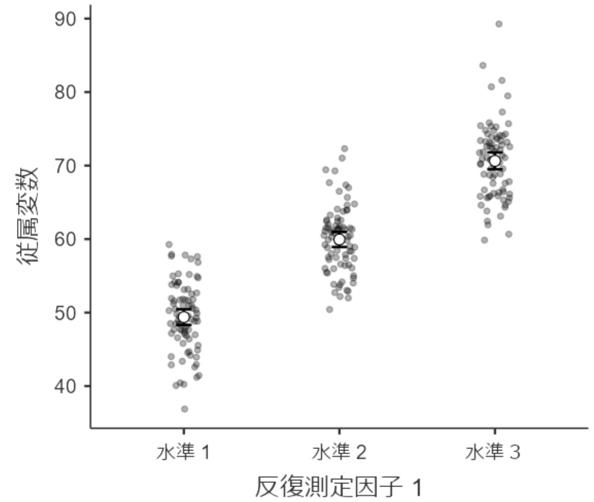


反復測定の分散分析　jamovi

反復測定の分散分析 jamovi

推定周辺平均

反復測定因子 1



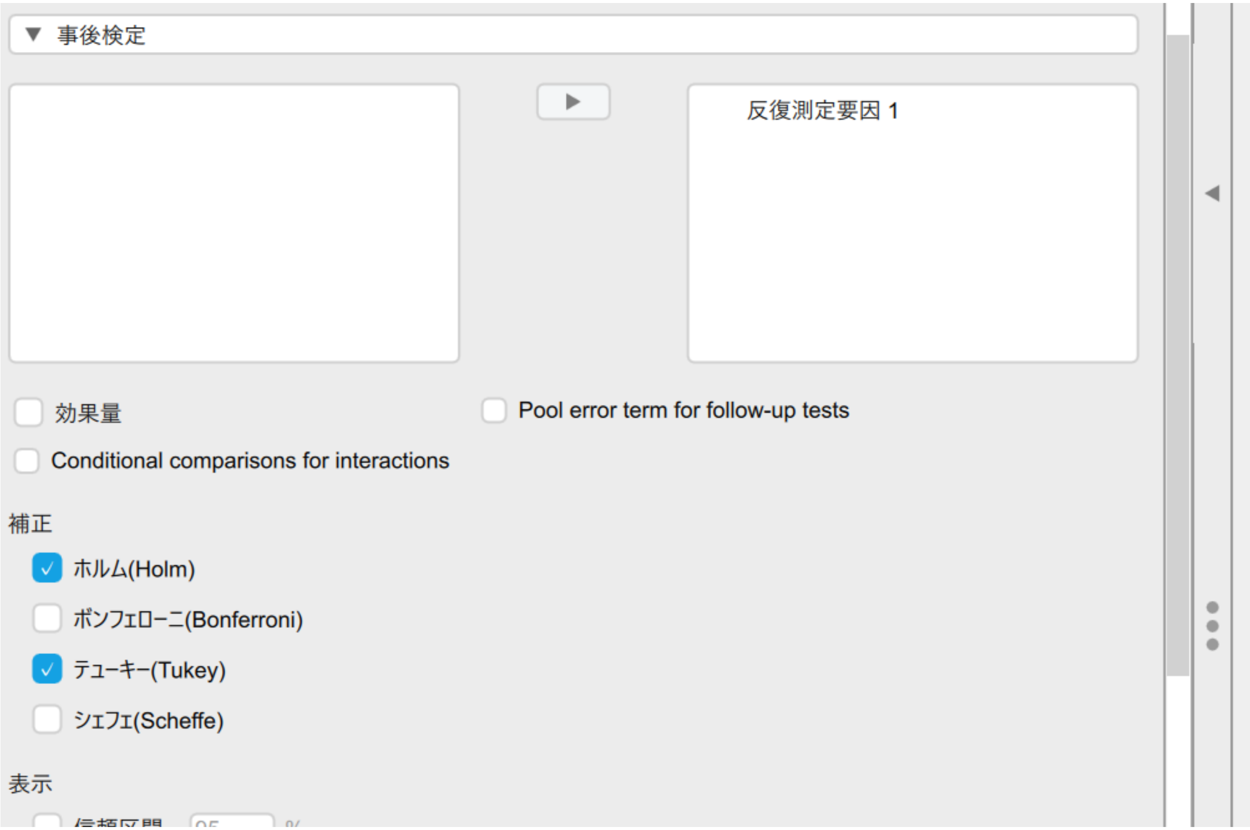
推定周辺平均値 - 反復測定因子 1

			95%信頼区間	
反復測定因子 1	平均値	標準誤差	下限	上限
水準 1	49.4	0.535	48.3	50.4
水準 2	59.9	0.513	58.9	61.0
水準 3	70.6	0.576	69.5	71.8

多重比較

- 反復測定の多重比較はHolm（ホルム）を選択
  - Tukeyは独立群が前提なので不適
  - といいつつ、古い研究ではTukeyが使われていたりすることがあるので注意
  - JASPでは反復でTukeyは原則出力されない

多重比較 JASP



水準 1	80	49.38	4.788	0.535	0.097
水準 2	80	59.94	4.587	0.513	0.077
水準 3	80	70.64	5.154	0.576	0.073

記述統計量プロット

雨雲プロット

事後検定

事後比較- 反復測定要因 1

		平均値差	標準誤差	df	t	$p_{テューキー}$	$p_{ホルム}$
水準 1	水準 2	−10.56	0.746	79	−14.16	.	< .001
	水準 3	−21.26	0.836	79	−25.42	.	< .001
水準 2	水準 3	−10.70	0.829	79	−12.91	.	< .001

注 3 の推定値の族を比較する調整された p 値

注 Tukey 修正されたp値は，反復測定の事後検定には適していません（Maxwell、1980; Field、2012）。

事後の多重比較　jamovi

事後検定

多重比較 - 反復測定因子 1

比較							
反復測定因子 1		反復測定因子 1	平均値の差	標準誤差	自由度	t	$p_{ホルム}$
水準 1	-	水準 2	−10.6	0.746	79.0	−14.2	< .001
	-	水準 3	−21.3	0.836	79.0	−25.4	< .001
水準 2	-	水準 3	−10.7	0.829	79.0	−12.9	< .001

[4]

推定周辺平均

反復測定因子 1