

統計解析

－ 分散分析の基本概念と実践 －

ファイル情報 - v2

編集者 1 - 島田匠悟

編集者 2 - 芹澤尚舜

書類作成日: 2023/08/26

最終更新日: 2024/11/19

東京都立大学大学院
システムデザイン研究科 情報科学域
西内研究室 修士1年
芹澤尚舜

目次

1

分散分析の概要

– 二元配置分散分析を中心に，分散分析の概要を説明

2

多重比較の概要

– 分散分析の下位検定で用いられる多重比較の説明

3

二元配置分散分析の実施

– js-STAR XR+を用いて，二元配置分散分析を体験

4

js-STAR XR+ × R言語

– js-STAR XR+と統計ソフトRを使用した分析手法

§ 1

1

分散分析の概要

– 二元配置分散分析を中心に，分散分析の概要を説明

2

多重比較の概要

– 分散分析の下位検定で用いられる多重比較の説明

3

二元配置分散分析の実施

– js-STAR XR+を用いて，二元配置分散分析を体験

4

js-STAR XR+ × R言語

– js-STAR XR+と統計ソフトRを使用した分析手法

そもそも分散分析とは…？

分散分析 (ANOVA : analysis of variance)

- 異なるグループの平均値に有意な差があるか否かを検定するための統計手法.
- 独立した要因(因子)を含むデータから, 各要因(因子)における水準間の平均値の差を検定するための分析方法.
- 要因(因子)の数によって分散分析にはバリエーションが存在する.
 1. 一元配置分散分析 (One-way ANOVA)
1つの要因(因子)からなるデータから, その因子に含まれる水準間の平均値の差を検定.
 2. 二元配置分散分析 (Two-way ANOVA)
2つの要因(因子)からなるデータから, 各因子における水準間の平均値の差を検定.
 3. 多元配置分散分析
3つ以上の要因(因子)からなるデータを分析する手法

分散分析－用語解説：要因，因子，水準

- 要因・因子

- － データの値に変化を与える要素のことを要因という．また，要因の中でも母平均に差をもたらすと考えられることから，研究対象となる要因を因子という．ただし，場合によってはこれらの用語が区別されずに用いられることも．

- － 例えば，研究で実験を行う場合…

- 要因：年齢，性別，気分，疲労度，etc.

- 因子：実験条件(研究で解明したいもの)

- 水準

- － 1つの因子に含まれる項目(グループ)のこと．

表1.1: 光量および肥料別の作物Aの収穫量データ(kg)

		光量		
		強	中	弱
肥料の種類	有機肥料	11.2	12.2	14.2
		10.7	11.5	15.4
	無機肥料	14.7	13.4	15.8
		13.3	14.7	16.3

赤枠：要因(因子)

緑枠：水準

分散分析とt検定

- どちらも母平均の検定を行う点では一緒であるが、t検定は1群、2群のデータに適用できるのに対し、分散分析は3群以上の際に適用可能な検定。
- 【分散分析】と呼ばれるのは、分散を使用して母平均の検定を行っているから。
- データの誤差(ばらつき)が要因(因子)によるものか、実験誤差によるものかを検定。
- 要因(因子)による誤差の方が大きい場合、母平均に差があるとする検定。
- データの差が、要因(因子)の違いによるものなのか、単なる偶然で生じたものなのか、どちらの確率が高いかを統計を使って計算する。

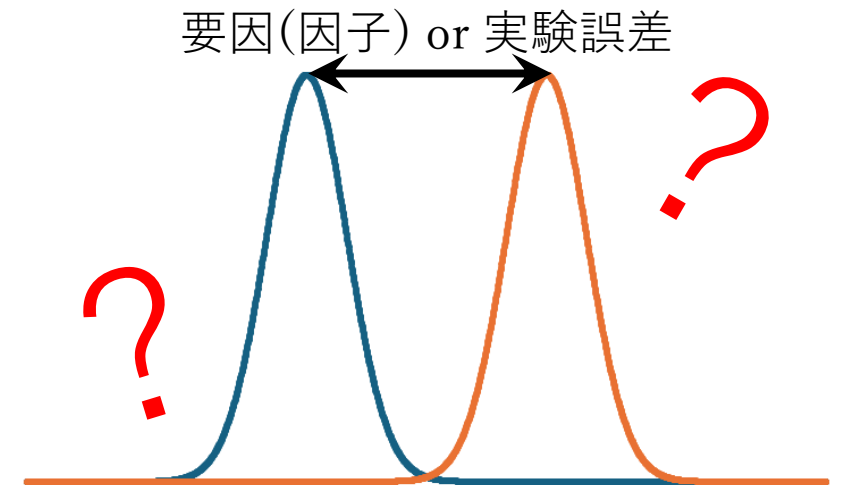


図1.1：分散分析のイメージ図

対象となるデータの性質

- パラメトリック検定 vs ノンパラメトリック検定
 - ー パラメトリック検定は解析するデータが特定の分布に従っていると仮定して行われる検定手法であるのに対し、ノンパラメトリック検定は解析データの分布に特定の分布を仮定せず(分布に依存せず)に用いることのできる検定手法.
 - ー 分散分析はパラメトリック検定に分類される.
- 分散分析にかけるデータには、次の二つの分布仮定が満たされる必要がある.
 1. 母集団が正規分布に従うこと.
 2. 等分散性が保たれていること.

(参考)データの尺度

- 名義尺度

- データがカテゴリに分類される尺度。単にグループを識別するためだけのもの。

- Ex: 性別(男性, 女性), 血液型(A型, B型, O型, AB型), 国名, etc.

- 順序尺度

- データに順序があるものの, 各々の順序の差が不明で, 具体的な間隔が不明な尺度。

- Ex: 満足度(1: 不満足, 2: 普通, 3: 満足), 教育水準(1: 高校卒, 2: 大学卒, 3: 大学院卒), etc.

- 間隔尺度 (分散分析などのパラメトリック検定に適したデータ尺度)

- 順序尺度に加え, 数字の間に等しい距離がある尺度。一方でゼロが絶対的でなく比率の比較は不可能。

- Ex: 日付, 温度, etc.

- 2°Cと7°C – 4°Cと9°Cは等しい変化量であるが, 40°Cは20°Cの2倍の暑さなどの比率的な意味は持たない。

- 比例尺度 (分散分析などのパラメトリック検定に適したデータ尺度)

- 間隔尺度に加え, 絶対的なゼロ点が存在する尺度。比率の比較は不可能。

- Ex: 重量, 距離, 身長, 年齢, 売上高, etc.

分散分析－用語解説：交互作用

交互作用とは…

- ある要因の水準間の平均値の差が、別の要因の水準によって異なること。
- 二つの要因が組み合わさることで初めて現れる**相乗効果**。

交互作用の具体例

- 分析：二元配置分散分析
 - 要因：要因A（肥料：有機，無機）と要因B（光量：強，弱）
 - 効果：ある作物の収穫量において、**要因A(肥料)が要因B(光量)の水準によって異なる影響を及ぼす**.
 - ・ 有機肥料を使用した場合，光量が強い条件下で収穫量が増加するが，弱い条件下で収穫量は減少。
 - ・ 無機肥料ではどちらの光量でも一定の効果を示すが，光量による影響が確認されない。
- このような関係性が確認される場合，肥料と光量の間に交互作用があることを示す。
- この交互作用を**一次の交互作用(A×B)**ともいう(二つの要因の組み合わせが結果に影響を及ぼす場合)。

分散分析－用語解説：主効果

主効果とは…

- ある一つの要因が水準間で従属変数に影響を与える効果のこと.
- 他の要因の水準や交互作用を一切考慮せず，一つの要因が結果にどのように影響するかを示す.

主効果の具体例

- 分析：二元配置分散分析
 - 要因：要因A（肥料：有機，無機）と要因B（光量：強，弱）
 - 効果：ある作物の収穫量において，**要因A(肥料)が要因B(光量)の水準によらず異なる影響を及ぼす.**
 - ・ 有機肥料を使用した場合において，光量に影響を受けず収穫量は増加.
 - ・ 一方で無機肥料を使用した場合において，光量に影響を受けず収穫量は減少.
- このような関係性が確認される場合，肥料に主効果があることを示す.

分散分析－用語解説：単純主効果

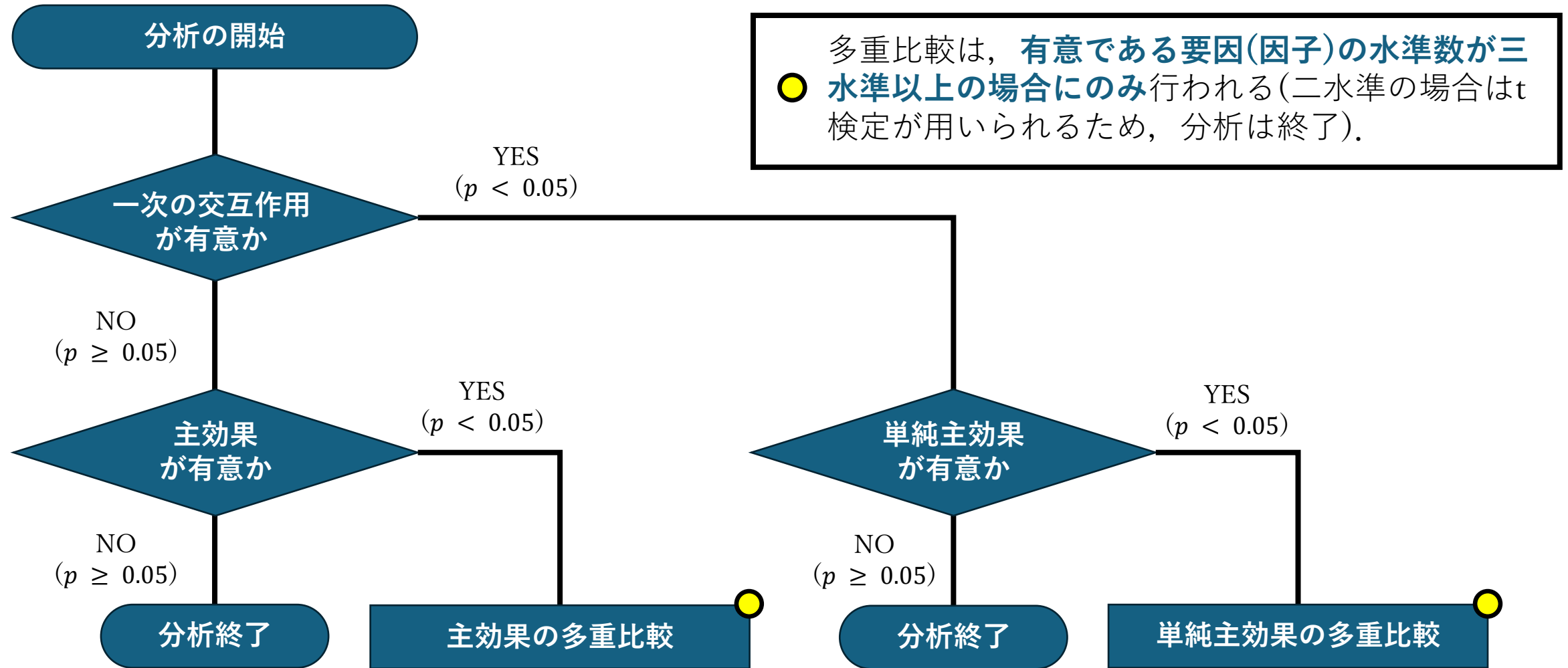
単純主効果とは…

- 二つの要因について、一方の要因の特定の水準における他の要因の主効果のこと.
- 一次の交互作用の下位検定として行われる(交互作用の詳細な理解のため).

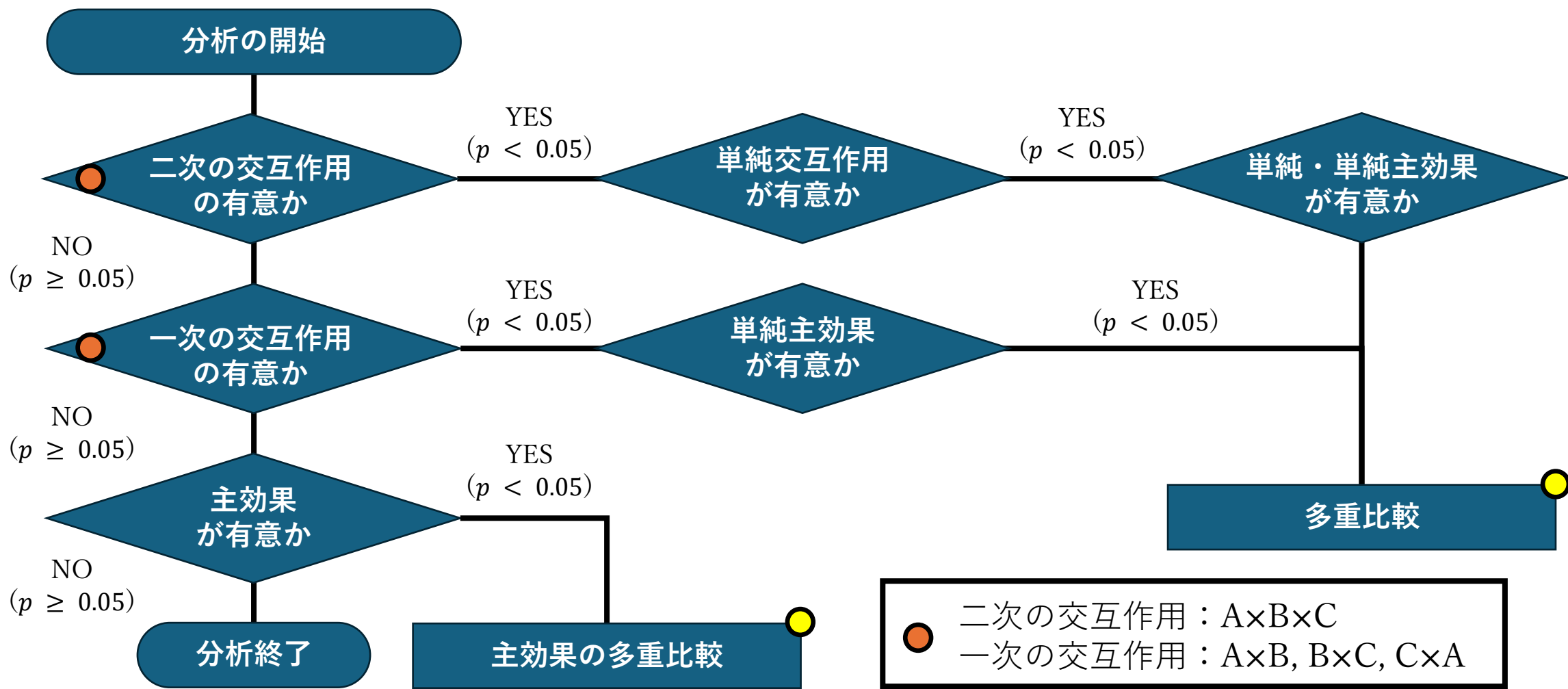
単純主効果の具体例

- 分析：二元配置分散分析
 - 要因：要因A（肥料：有機，無機）と要因B（光量：強，弱）
 - 効果：ある作物の収穫量において，**要因A(肥料)が要因B(光量)の特定の水準において異なる影響を及ぼす.**
 - ・ 光量が強い条件下で(要因Bの水準を固定して)，有機肥料と無機肥料の収穫量を比較する.
 - ・ 有機肥料を使用すると，無機肥料を使用した場合と比べ収穫量が増加することが確認.
- このような関係性が確認される場合，光量が強い条件下において肥料の単純主効果があることを示す.

二元配置分散分析のフロー



(参考)三元配置分散分析のフロー



§ 2

1

分散分析の概要

– 二元配置分散分析を中心に，分散分析の概要を説明

2

多重比較の概要

– 分散分析の下位検定で用いられる多重比較の説明

3

二元配置分散分析の実施

– js-STAR XR+を用いて，二元配置分散分析を体験

4

js-STAR XR+ × R言語

– js-STAR XR+と統計ソフトRを使用した分析手法

多重比較－用語解説

- 帰無仮説

- ある仮説が正しいかどうかの判断のために立てられる仮説。たいていは否定されることを期待して立てられる。帰無仮説と真逆の意味をもつ仮説を対立仮説という。

- 第1種の過誤

- 帰無仮説が正しいにもかかわらず、帰無仮説を誤って棄却すること。反対に対立仮説が正しいにもかかわらず、帰無仮説が棄却されずに誤って帰無仮説を支持することを第2種の過誤という。

- 有意水準

- 帰無仮説を設定したとき、その仮説を棄却するか否かを判定する基準となる確率のことで検定を行う前に設定されるもの。有意水準5%で検定を行うということは、第1種の過誤を犯す確率が5%であるということ。
 - 「すべての実験条件で平均値が等しい」とするのが帰無仮説。

多重比較の手法

多重比較とは…

- 水準間の平均値の差を比較する検定手法。 **水準の数が三つ以上の比較において、どの水準間に差があるかを調べる**ことができる。
- なお、js-STAR XR+で使用可能な多重比較手法は以下の3種類である。
 1. LSD (least significant difference) 法
第一種の過誤のリスクが高いことが特徴。
 2. Bonferroni法
第二種の過誤のリスクが高いことが特徴。
 3. Holm法
西内研では、この手法が主に用いられる。

→ p 値を調整し、検定全体の有意水準 α が0.05を超えないように働く。

表2.1: 光量および肥料別の作物Aの収穫量データ(kg)

		光量		
		強	中	弱
肥料の種類	有機肥料	11.2	12.2	14.2
		10.7	11.5	15.4
	無機肥料	14.7	13.4	15.8
		13.3	14.7	16.3

t検定の繰り返し使用におけるリスク

水準の数が三つの多重比較は，t検定で2群毎に独立した検定を3回行えばできる!?

→ **検定の多重性の問題が発生**

多重性の問題の具体例

- 3種類の治療法(A, B, C)をそれぞれ比較する．このとき検定を3回繰り返すが，1回の検定を5%の有意水準で行うため，最終的な有意水準が5%よりも大きくなる(第一種の過誤を犯す確立が上昇する)．

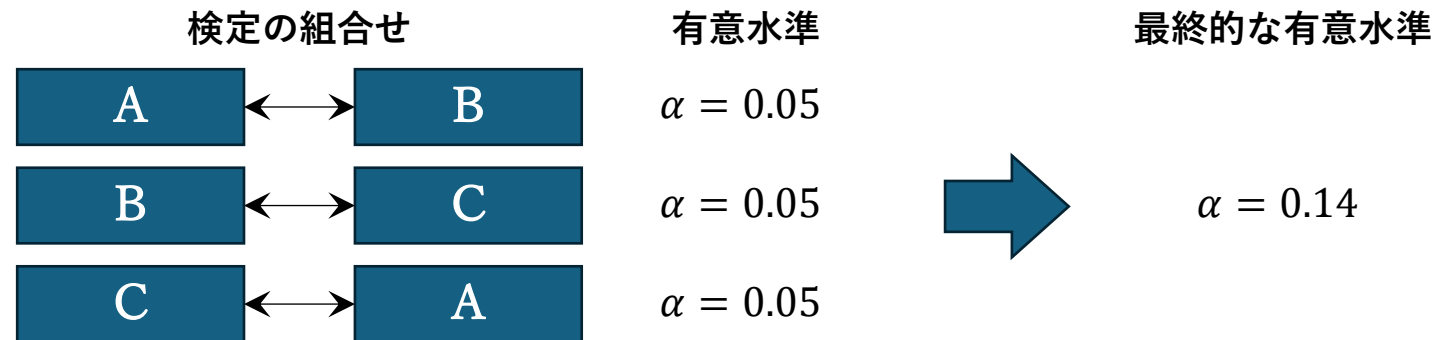


図2.1：多重性の問題のイメージ図

- 多重比較法は，このような問題を回避するために，全体の有意水準をあらかじめ設定している有意水準になるよう調整を行ってくれる．

多重性の問題 – 数学的な理解

- 水準の数が三つの多重比較を， t検定を用いて2群毎に有意水準0.5%で3回の独立した検定を行った場合において， 検定全体の第一種の過誤率を求める.
- 求める第一種の過誤率は**少なくとも1回の検定で帰無仮説を棄却する確率**と同値.

— 1回の検定で帰無仮説を棄却しない確率

$$1 - 0.05 = 0.95$$

— 独立した検定を3回繰り返す場合， すべての検定で帰無仮説を棄却しない確率

$$0.95^3 = 0.857$$

— 少なくとも一回の検定で帰無仮説を誤って棄却する確率

$$1 - 0.95^3 = 1 - 0.857 = 0.143$$

- 以上より， t検定を用いて2群毎に3回の独立した検定を行う場合， 約14.3%の確立で帰無仮説を誤って棄却し， 検定全体の有意水準が5%を上回る.

多重比較 – part1 : LSD法(1/2)

- t統計量に基づいた手法で，4群以上で使用できないという制約がある．
- 有意水準(LSD法では必ず5%)をクリアする最小の幅を計算し，その幅が条件の差よりも大きい場合に有意，小さい場合に有意でないとする検定．

Step1 : LSD値を求める．

$$LSD = t\text{値} \times \sqrt{MS \times \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}$$

(t値は有意水準5%のときの値)
(N_1, N_2 は各条件のデータ数)
(MS は偶然の差の平均平方)

Step2 : 各条件の平均の差とLSDを比較

case2.1 : $LSD <$ 各条件の差 … 有意差有り(帰無仮説を棄却し，対立仮説を支持)

case2.2 : $LSD >$ 各条件の差 … 有意差なし(帰無仮説は棄却されない)

多重比較 – part1 : LSD法(2/2)

偶然の差の平均平方(MS)の導出式

$$\begin{aligned} MS = & (A_1 - \overline{X_A})^2 + (A_2 - \overline{X_A})^2 + (A_N - \overline{X_A})^2 \\ & + (B_1 - \overline{X_B})^2 + (B_2 - \overline{X_B})^2 + (B_N - \overline{X_B})^2 \\ & + (C_1 - \overline{X_C})^2 + (C_2 - \overline{X_C})^2 + (C_N - \overline{X_C})^2 \end{aligned}$$

$\overline{X_A}$: 条件Aの平均

$\overline{X_B}$: 条件Bの平均

$\overline{X_C}$: 条件Cの平均

表2.2 : LSD法説明のためのデータ概要

	A	B	C
1	A_1	B_1	C_1
2	A_2	B_2	C_2
\vdots	\vdots	\vdots	\vdots
N	A_N	B_N	C_N

多重比較 – part2 : Bonferroni法

- 検定を1つに絞れず，かつ検定の優先順位をつけることができない場合，有意水準を調整し，多重性の問題を解決する手法の一つ.

$$\alpha_N = \frac{\alpha_S}{N}$$

α_N : 検定一つ当たりの有意水準

α_S : 基準となる(通常の)有意水準(多くの場合 0.05)

N : 検定総数

- p 値が α_N を下回れば有意差があるとする検定.
- ただし，**他の有意水準の調整方法より基準が厳しく，有意になりにくい傾向.**
 - 第2種の過誤のリスクが高い.
 - より厳密な検定を行うことが可能.

多重比較 – part3 : Holm法

- Bonferroni法を改良した手法.
- 有意水準の調整手法がBonferroni法と異なり，検定の厳密さが緩くなっているため，より有意差が出やすい.
- p 値の大きさに従って，有意水準が異なることが特徴.

Step1 : N 個の帰無仮説を p 値の小さい順に並べる.

Step2 : 最も p 値が小さい第1順位の帰無仮説の有意水準を α/N に設定する.

case 2.1 : p 値 $< \alpha/N$ ならば第1順位の帰無仮説を棄却し，対立仮説を採択.

case 2.2 : p 値 $> \alpha/N$ ならば第1順位以下の全ての帰無仮説の判定を保留とする.

Step3 : もし第1順位の帰無仮説が棄却された場合，第2順位の帰無仮説の有意水準を $\alpha/(N-1)$ に設定する.

case 3.1 : p 値 $< \alpha/(N-1)$ ならば第2順位の帰無仮説を棄却.

case 3.2 : p 値 $> \alpha/(N-1)$ ならば第2順位以下の全ての帰無仮説の棄却を保留とする.

Step4 : 上記の手順を繰り返し行う.

case 4.1 : 第 $(k-1)$ 順位の帰無仮説が棄却された場合，第 k 順位の有意水準を $\alpha/(N+1-k)$ に設定する.

§ 3

1

分散分析の概要

– 二元配置分散分析を中心に，分散分析の概要を説明

2

多重比較の概要

– 分散分析の下位検定で用いられる多重比較の説明

3

二元配置分散分析の実施

– js-STAR XR+を用いて，二元配置分散分析を体験

4

js-STAR XR+ × R言語

– js-STAR XR+と統計ソフトRを使用した分析手法

js-STAR XR+とは…？

- オンラインで素早くデータ分析ができる無償の統計ソフト([リンク](https://www.kisnet.or.jp/nappa/software/star/index.htm)).
- t検定, 分散分析, 多変量解析, ノンパラメトリック法など多種多様な分析が可能.
- ダウンロードして, オフラインで動作させることも可能(グラフ表示などの一部機能はネットに接続されている必要あり).



図3.1：js-STAR XR+のwebサイトのスクリーンショット

実験計画法(1/3)

表3.1：実験計画法と分散分析の略号の対応関係

実験計画法		分散分析の略号
1要因(因子)	参加者間計画	AS
	参加者内計画	SA
2要因(因子)	参加者間計画	ABS
	参加者内計画	SAB
	混合計画	ASB
3要因(因子)	参加者間計画	ABCS
	参加者内計画	SABC
	混合計画 I	ABSC
	混合計画 II	ASBC

S:参加者
ABC: 独立の1要因

分散分析

- A s (1 要因参加者間)
- s A (1 要因参加者内)
- A B s (2 要因参加者間)
- A s B (2 要因混合)
- s A B (2 要因参加者内)
- A B C s (3 要因参加者間)
- A B s C (3 要因混合)
- A s B C (3 要因混合)
- s A B C (3 要因参加者内)

図3.2：js-STAR XR+のwebサイトの分散分析

統計ソフトを使用する前に、実施した実験の実験計画法及びそれと対応する検定手法を把握，理解しておく必要がある。

実験計画法(2/3)

- 参加者間計画(対応無し, between-subjects design)
 - 水準間で参加者が異なる実験のこと.
 - 各参加者から1つの条件でしかデータを得ない。
- 参加者内計画(対応あり, within-subjects design)
 - 要因内の全ての水準で, 各参加者が実験に参加.
 - つまり, 各参加者から全ての条件でデータを得る.
- 混合計画(Mixed Design)
 - 参加者間計画と参加者内計画の両方を含む実験.

実験計画法(3/3)

表3.2：2要因 – 2×3参加者間計画

A1			A2		
B1	B2	B3	B1	B2	B3
田中	阿部	鈴木	橋口	堀井	倉沢
山際	浜本	吉川	染谷	植竹	西
遠藤	金森	岡崎	野瀬	高野	小島

参加者は**どれか1つの条件に割り当て**(実験参加者数 $N = 18$)

表3.3：2要因 – 2×3参加内間計画

A1			A2		
B1	B2	B3	B1	B2	B3
田中	田中	田中	田中	田中	田中
山際	山際	山際	山際	山際	山際
遠藤	遠藤	遠藤	遠藤	遠藤	遠藤

各参加者は**すべての条件に割り当て**(実験参加者数 $N = 3$)

表3.4：2要因 – 2×3混合計画

A1			A2		
B1	B2	B3	B1	B2	B3
田中	田中	田中	阿部	阿部	阿部
山際	山際	山際	浜本	浜本	浜本
遠藤	遠藤	遠藤	金森	金森	金森

要因Aは参加者間，**要因Bは参加者内**(実験参加者数 $N = 6$)

js-STAR XR+の操作方法 - 2要因参加者内

データ入力欄

- データを各々の欄に入力することが可能.
- ただし、データ総数が増加した場合の手入力は大変...
Ex: 2要因3水準の実験を20名に行った場合、データ総数は180.
- 緑枠をクリックすると、下の図3.3のスペースが展開. ここにはエクセルに保存されているデータのコピー&ペーストが可能.
- コピー後は図3.3の緑下線の【閉】が【代入】に遷移するので、**【代入】をクリックしデータを入力.**

参加者	A1		A2	
	B1	B2	B1	B2
1				
2				

☐ LSD法 ☐ HSD法 ☐ Bonferroni法 ☒ Holm法

図3.3: 図3.4の緑枠クリック後のUI

s A B デザイン (2 要因参加者内)

メイン データ形式 グラフ 説明

データ

読み込み 保存 消去 シミュレーション

参加者数: 2

要因A 名_前: A ×
要因A 水準数: 2

要因B 名_前: B ×
要因B 水準数: 2

参加者	A1		A2	
	B1	B2	B1	B2
1				
2				

☐ LSD法 ☐ HSD法 ☐ Bonferroni法 ☒ Holm法

● Rオプション ●

多重比較のp値調整法: BH法

☐ ベイズファクタ: 【パッケージ BayesFactor が必要】

計算!

検定開始ボタン

図3.4: sABのUI

js-STAR XR+で分散分析を実施

2要因参加者内計画での実験を仮定して，作物Aの収穫量データで分散分析を実施.

- 要因A – 3水準：肥料(有機肥料，化学肥料，無機肥料)
- 要因B – 3水準：光量(強，中，弱)
- 従属変数：作物Aの収穫量(kg)

Step1：エクセルをダウンロード(リンク).

Step2：Webでjs-STARを開き，検定のデータ形式を設定.

参加者数：50

要因A：肥料 – 水準数 3

要因B：光量 – 水準数 3

Step3：エクセルを開き， C5:K54の範囲でデータをコピー.

Step4：前頁のスライドを参照し，データをペーストし入力.

Step5：Holm法にチェックが入っていることを確認.

Step6：検定開始のため【計算！】ボタンを押下.

The screenshot shows the 's A B デザイン (2 要因参加者内)' screen. It has a top navigation bar with 'メイン', 'データ形式', 'グラフ', and '説明'. The 'データ形式' tab is selected. Below the navigation bar, there is a 'データ' section with '読み込み', '保存', '消去', and 'シミュレーション' buttons. The 'シミュレーション' button is highlighted. The main area contains input fields for '参加者数' (50), '要因A 名_前' (肥料), '要因A 水準数' (3), '要因B 名_前' (光量), and '要因B 水準数' (3). Each input field has a small 'x' button to the right.

図3.5：検定のパラメータ設定のスクリーンショット

js-STAR XR+の出力結果の見方

== Analysis of Variance ==				
A(3) = 肥料 B(3) = 光量				
S. V	SS	df	MS	F
subj	1068.7297	49	21.8108	
A	517.2506	2	258.6253	12.51 **
sxA	2025.5261	98	20.6686	
B	5.9357	2	2.9679	0.13 ns
sxB	2234.3553	98	22.7995	
AxB	1198.9752	4	299.7438	13.94 **
sxAxB	4214.2474	196	21.5013	
Total	11265.0201	449	+p<.10 *p<.05 **p<.01	

図3.6：分散分析の結果

== Analysis of AxB Interaction ==				
S. V	SS	df	MS	F
A at B1: (sxA at B1:	100.1307 1673.3718	2 98	50.0654 17.0752)	2.93 +
A at B2: (sxA at B2:	101.7849 2154.5785	2 98	50.8924 21.9855)	2.31 ns
A at B3: (sxA at B3:	1514.3101 2411.8233	2 98	757.1551 24.6104)	30.77 **

図3.7：一次の交互作用の分析結果(一部)

→ 有意傾向あり ($p < 0.1$)

→ 有意差なし ($p \geq 0.1$)

→ 有意差あり ($p < 0.01$)

交互作用A×Bが有意である($p < 0.01$)

→ 有意水準1%で「実験結果に肥料と光量の交互作用が確認された」という主張が成り立つ。

多重比較の結果の解釈の仕方

B(光量) at A1 (肥料：有機肥料)の単純主効果の多重比較を例に挙げる.

- 「B1(光量：強)」と「B2 (光量：中)」には有意差($p < 0.05$)が見られる.
- 「B2(光量：中)」と「B3 (光量：弱)」には有意差($p < 0.05$)が見られる.
- 「B2(光量：中)」と「B3 (光量：弱)」には有意差($p < 0.05$)が見られる.

→ 有意水準5%で，光量の違いが有機肥料を使用した際の作物Aの収穫量に有意差をもたらすことが確認された.

== Multiple Comparisons by Holm ==				
B at A1 Level				
(MSe= 19.8514, * p<.05)				
B1	<	B2	*	(alpha' = 0.0500)
B1	<	B3	*	(alpha' = 0.0167)
B2	<	B3	*	(alpha' = 0.0250)

図3.8： B(光量) at A1 (肥料：有機肥料)の単純主効果の結果

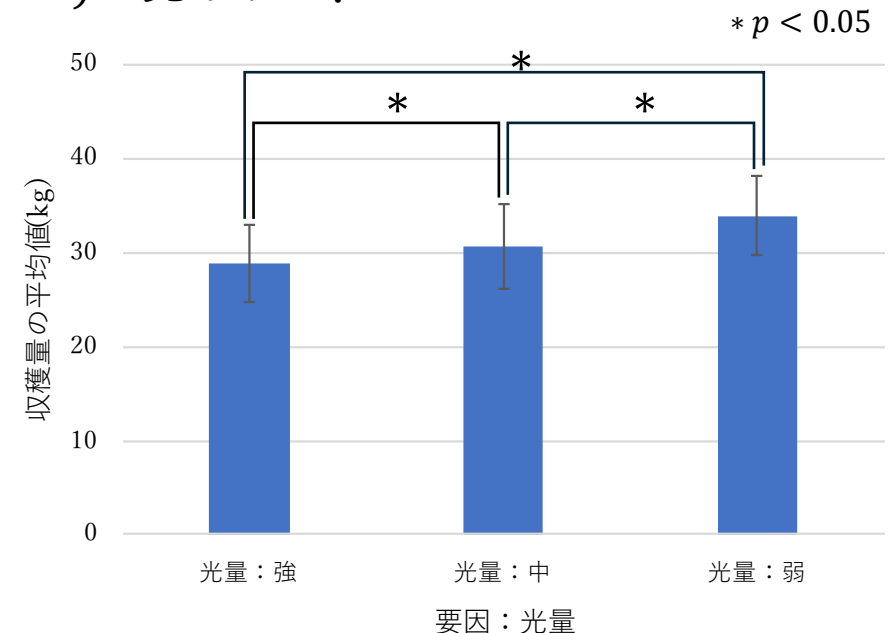


図3.9： 肥料－有機肥料条件下での光量別作物Aの収穫量(kg)

統計的に有意差があるとは

「光量の違いが、有機肥料における収穫量に有意差をもたらす」とは…？

- **そもそも「統計的に有意な差がある」とは**，平均値が群間で異なっているということの意味するわけではなく，**群間での平均値が等しいとは言えないという消極的な意味**のこと。
- つまり今回の場合，有機肥料における収穫量について，
 - ✗ 光量－三水準間での平均値が異なる。
 - 光量－三水準間での平均値が等しいとは言えない。
- 100回中95回(有意水準5%と同値)は「有機肥料における収穫量について，光量－三水準間での平均値が等しいとは言えない」といえるため，おそらく偶然ではないのでとりあえずこちらを正しいとしましょうという意味。

分散分析の交互作用－図からの確認

交互作用とは…

ある要因の水準間の平均値の差が、別の要因の水準によって異なること。

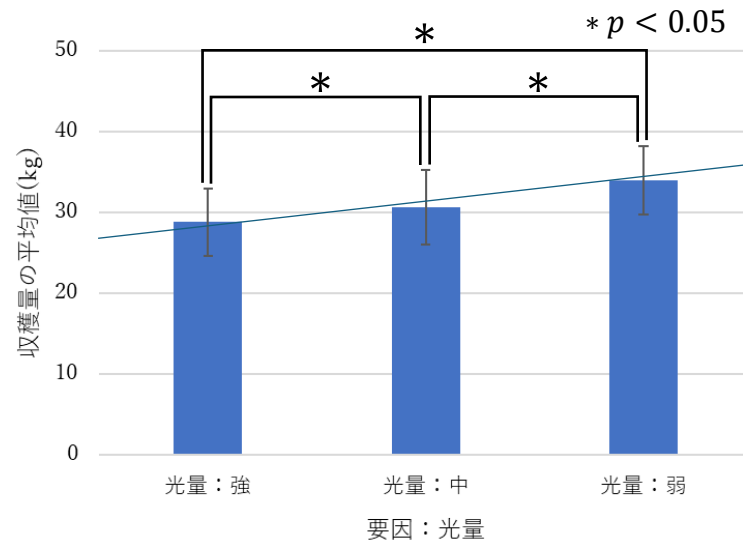


図3.10： 肥料－有機肥料条件下での光量別作物Aの収穫量(kg)

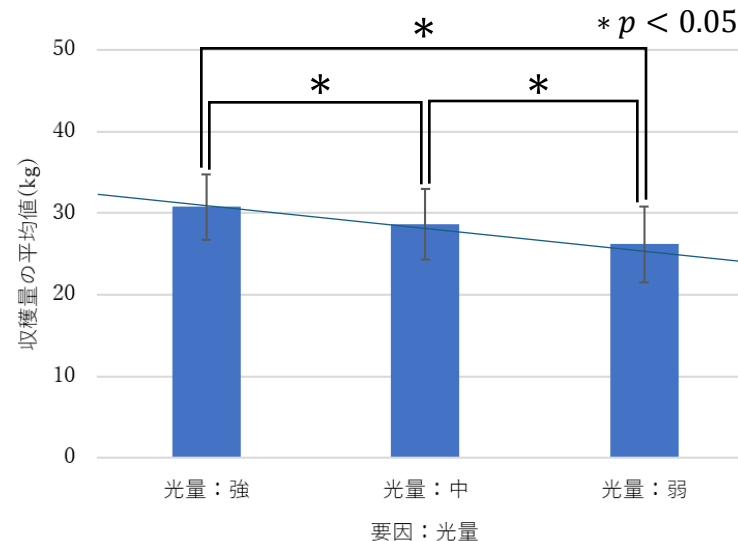


図3.11： 肥料－化学肥料条件下での光量別作物Aの収穫量(kg)

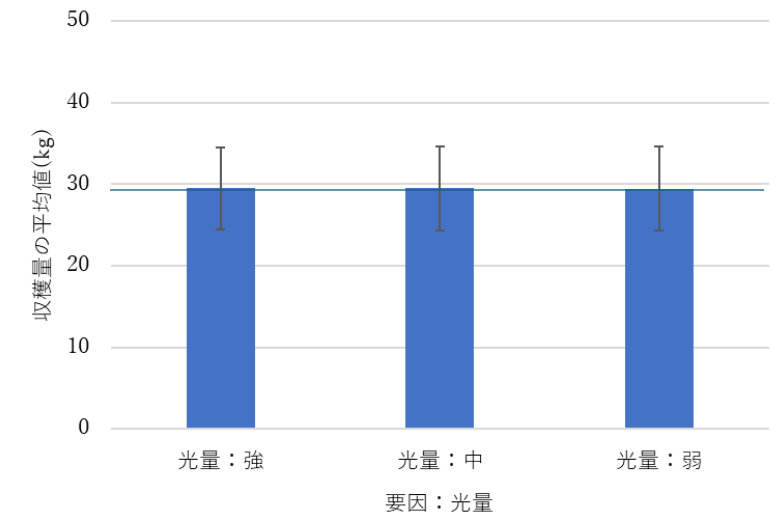


図3.12： 肥料－無機肥料条件下での光量別作物Aの収穫量(kg)

§4

1

分散分析の概要

– 二元配置分散分析を中心に，分散分析の概要を説明

2

多重比較の概要

– 分散分析の下位検定で用いられる多重比較の説明

3

二元配置分散分析の実施

– js-STAR XR+を用いて，二元配置分散分析を体験

4

js-STAR XR+ × R言語

– js-STAR XR+と統計ソフトRを使用した分析手法

js-STAR XR+ × R言語

- js-STARにはR言語のオプションがあり，R用のコードが自動的に生成されるようになっている．
- より詳細な分析やその結果の書き方を参考にすることができるため，こちらでのやり方を推奨．
- 以下ではRで実行するための環境構築から実行までを手順ごとに説明する．

Step1：R言語の環境構築．

環境はRStudioを使用．環境構築自体はインストールするのみなので簡単．

Step2：js-STAR XR+での操作．

R用のオプションを設定する必要あり．

Step3：RStudioでの実行．

パッケージのインストール，実行方法，実行結果の確認の仕方について説明．

Step1：R言語の環境構築

以下のリンクを踏み，図4.1に従ってRとOSに合ったRStudioをインストール。

リンク：<https://posit.co/download/rstudio-desktop/>

The screenshot shows the RStudio download page. On the left, under '1: Install R', there is a button 'DOWNLOAD AND INSTALL R' highlighted with a red box. An arrow points from this button to the '2: Install RStudio' section. In the '2: Install RStudio' section, there is a button 'DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS' highlighted with a red box, with an arrow pointing to it from the '1: Install R' section. Below this button, it says 'Size: 263.71 MB | SHA-256: 2C3CF96A | Version: 2024.09.1+394 | Released: 2024-11-04'. To the right of this is the text 'Windows用 RStudio'. On the right side of the screenshot, there is a table with columns 'OS', 'Download', 'Size', and 'SHA-256'. The table lists download links for Windows 10/11, macOS 12+, Ubuntu 20/Debian 11, and Ubuntu 22/Debian 12. The macOS 12+ row has a button 'RSTUDIO-2024.09.1-394.DMG' highlighted with a red box, with an arrow pointing to it from the 'Windows用 RStudio' text. To the right of this button is the text 'Mac用 RStudio'.

OS	Download	Size	SHA-256
Windows 10/11	RSTUDIO-2024.09.1-394.EXE		2C3CF96A
macOS 12+	RSTUDIO-2024.09.1-394.DMG	613.31 MB	F1AAC1C8
Ubuntu 20/Debian 11	RSTUDIO-2024.09.1-394-AMD64.DEB	202.43 MB	2890EDE2
Ubuntu 22/Debian 12	RSTUDIO-2024.09.1-394-AMD64.DEB	202.43 MB	94896D5D

(a)：RとWindows用 RStudioのインストール

(a)：Mac用 RStudioのインストール

Rのインストール

図4.1：Rの環境構築のためのWebサイトスクリーンショット

Step2 : js-STAR XR+での操作(1/2)

Step2.1 : 検定のデータ形式を設定

s A B デザイン (2 要因参加者内)

メイン データ形式 グラフ 説明

データ

読込 保存 消去 シミュレーション

参加者数 : 50

要因A 名_前 : 肥料 ×

要因A 水準数 : 3

要因B 名_前 : 光量 ×

要因B 水準数 : 3

図4.2 : js-STARで検定のデータ形式設定後の状態

Step2.2 : データ入力

29.37	30.56	35.67	27.78	21.04	29.55	37.06	29.21
26.57	42.06	33.29	22.08	25.38	30.87	33.76	16.71
34.29	32.04	29.99	22.64	26.58	29.11	26.38	31.41
22.52	32.99	30.36	27.79	29.21	30.41	29.57	32.91
32.43	39.31	24.89	31.35	29.51	33.14	30.24	21.21
23.46	28.03	30.04	33.16	33.13	29.99	23.49	25.01
28.23	27.97	35.62	26.61	21.06	35.49	27.53	37.31
26.53	36.63	36.31	29.95	19.95	28.63	24.99	23.41
24.57	30.51	30.85	28.87	28.55	24.72	27.01	23.51

代入

☐ LSD法 ☐ HSD法 ☐ Bonferroni法 ☒ Holm法

● Rオプション ●

多重比較のp値調整法 : Holm法

☐ バイズファクタ : 【パッケージ BayesFactor が必要】

計算!

図4.3 : js-STARで検定のデータ入力後の状態

Step2 : js-STAR XR+での操作(2/2)

Step2.3 : Rオプションの設定と検定開始

29.37	30.56	35.67	27.78	21.04	29.55	37.06	29.2
26.57	42.06	33.29	22.08	25.38	30.87	33.76	16.7
34.29	32.04	29.99	22.64	26.58	29.1	26.38	31.4
22.52	32.99	30.36	27.79	29.2	30.41	29.57	32.9
32.43	39.31	24.89	31.35	29.5	33.14	30.24	21.2
23.46	28.03	30.04	33.16	33.13	29.99	23.49	25.0
28.23	27.97	35.62	26.6	21.06	35.49	27.53	37.3
26.53	36.63	36.3	29.95	19.95	28.63	24.99	23.4
24.57	30.5	30.85	28.87	28.55	24.72	27.01	23.5

代入

☐ LSD法 ☐ HSD法 ☐ Bonferroni法 ☒ Holm法

● Rオプション ●

多重比較のp値調整法: Holm法

☐ ベイズファクタ: 【パッケージ BayesFactor が必要】

設定後, 計算を押下 計算!

図4.4 : js-STAR XR+でRオプションでの多重比較のp値調整法をHolm法に設定後の状態

Step2.4 : Rプログラムの第一枠のコピー

Rプログラム

第一枠

保存 コピー 削除 伸▼ ▲縮

```
##### 分散分析 s A B (2 要因参加者内) : 主分析
# パッケージ psych が必要

## js-STARからの入力
levA = 3          # level of A
levB = 3          # level of B
n = c( # データ数
  50, 50, 50,
  50, 50, 50,
  50, 50, 50
)
data = c(
  29.37, 30.56, 35.67, 27.78, 21.04, 29.55, 37.06, 29.2,
  26.57, 42.06, 33.29, 22.08, 25.38, 30.87, 33.76, 16.7,
  34.29, 32.04, 29.99, 22.64, 26.58, 29.1, 26.38, 31.4,
  22.52, 32.99, 30.36, 27.79, 29.2, 30.41, 29.57, 32.9,
  32.43, 39.31, 24.89, 31.35, 29.5, 33.14, 30.24, 21.2,
  23.46, 28.03, 30.04, 33.16, 33.13, 29.99, 23.49, 25.0,
  28.23, 27.97, 35.62, 26.6, 21.06, 35.49, 27.53, 37.3,
  26.53, 36.63, 36.3, 29.95, 19.95, 28.63, 24.99, 23.4,
  24.57, 30.5, 30.85, 28.87, 28.55, 24.72, 27.01, 23.5
```

図4.5 : js-STAR XR+でのRプログラムの第一枠のコピー

Step3：RStudioでの実行(1/3)

Step3.1：R Scriptをエディタで展開

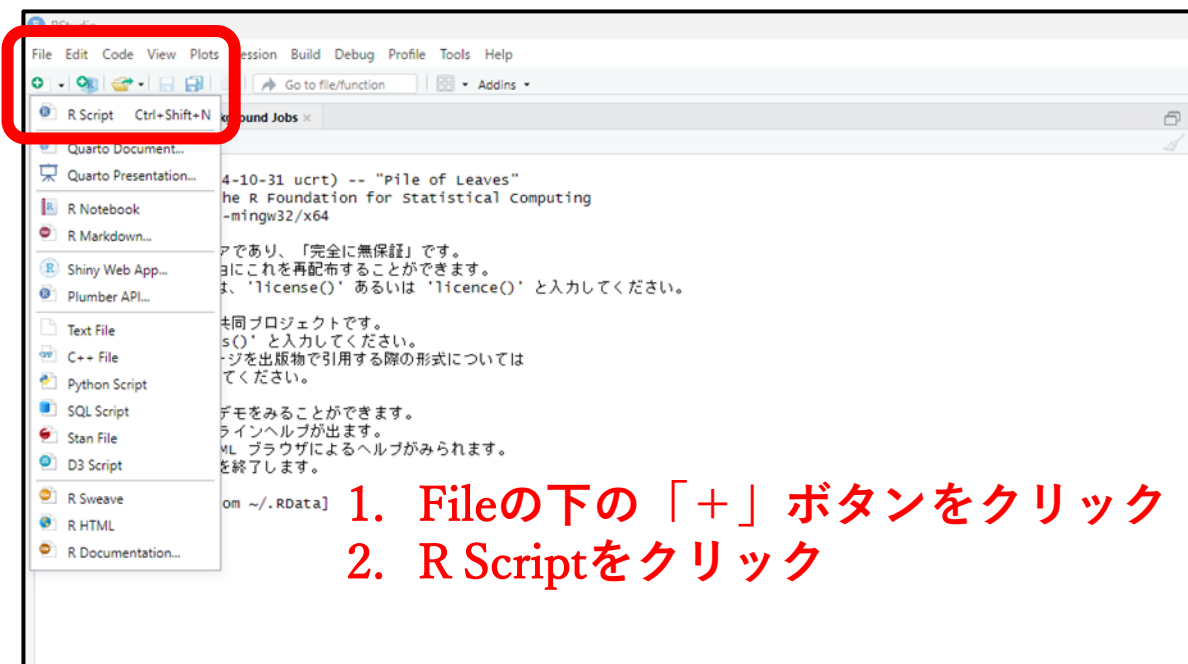


図4.6：RStudioでスクリプトを展開する手順

Step3.2：コピーした第一枠のペースト

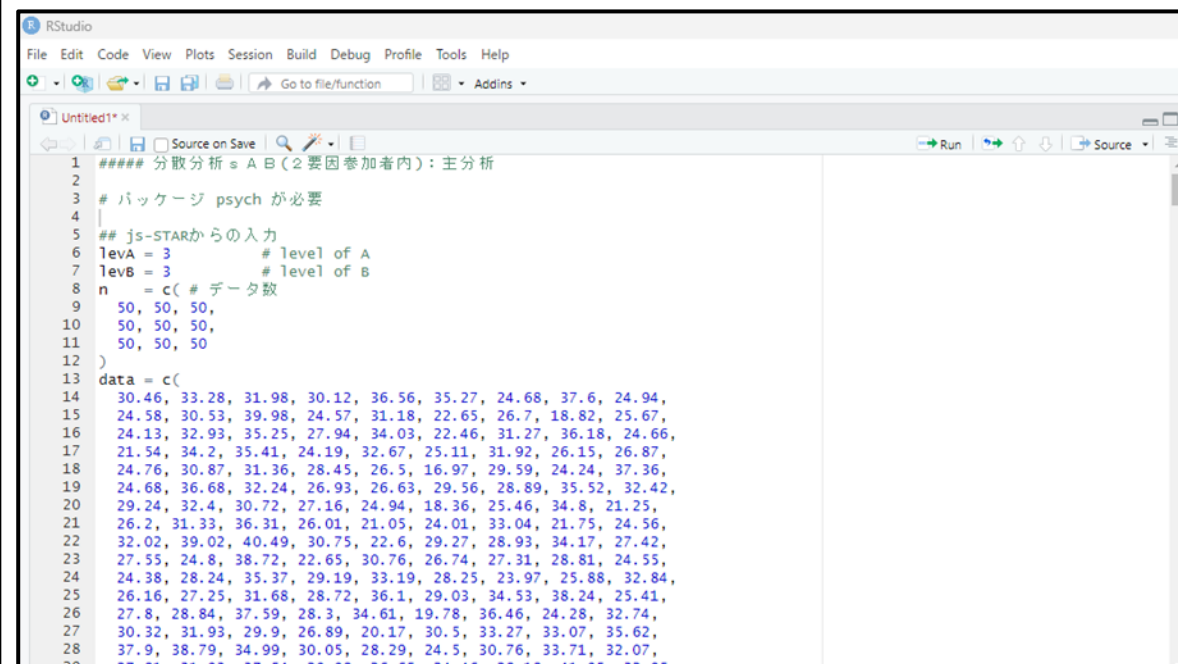


図4.7：Rプログラムの第一枠のペースト後の状態

Step3：RStudioでの実行(2/3)

Step3.3：パッケージのインストール※1

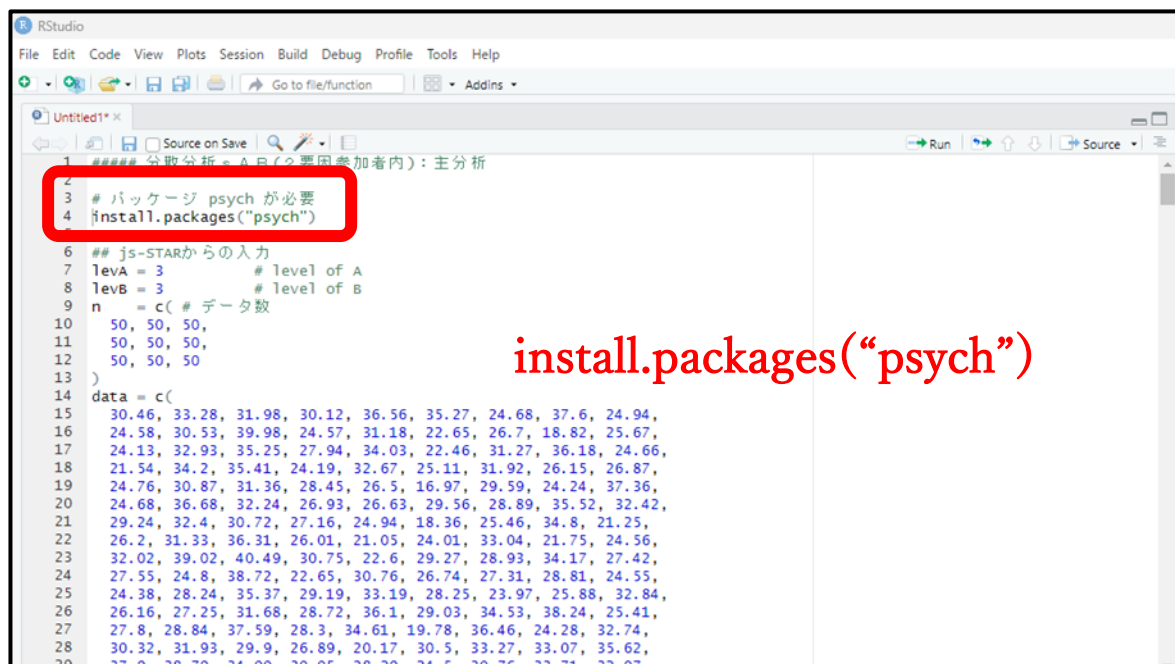


図4.8：RStudioでのパッケージのインストール

Step3.4：コードの実行

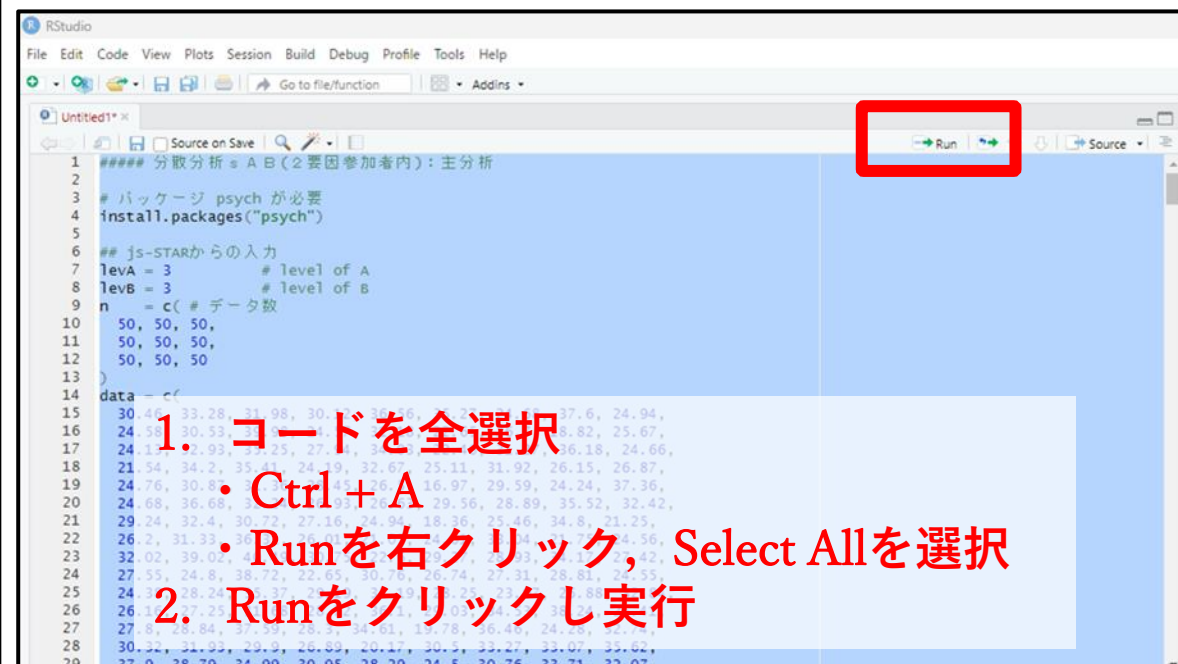
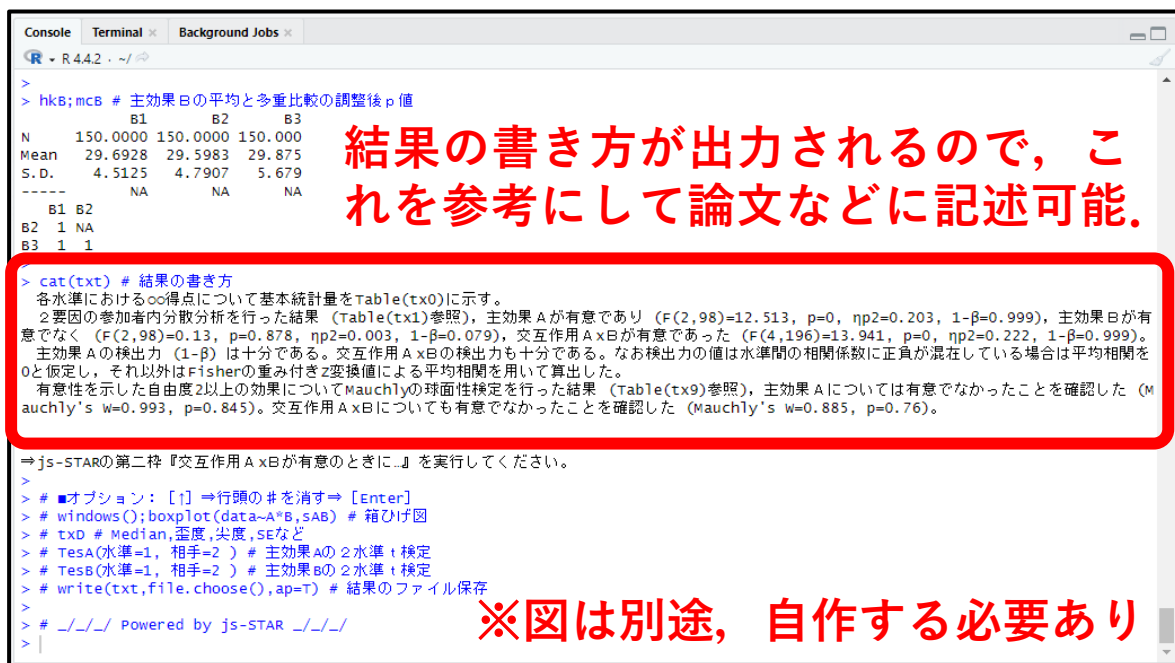


図4.9：Rプログラムの実行手順

※1：インストールは初回実行時のみ行う

Step3 : RStudioでの実行(3/3)

Step3.5 : Consoleで結果の確認



```
> hks;mcB # 主効果Bの平均と多重比較の調整後p値
      B1      B2      B3
N      150.0000 150.0000 150.000
Mean    29.6928 29.5983 29.875
S.D.     4.5125 4.7907 5.679
-----
      NA      NA      NA
B1 B2
B2 1 NA
B3 1 1

> cat(txt) # 結果の書き方
各水準における平均点について基本統計量をTable(tx0)に示す。
2要因の参加者内分散分析を行った結果 (Table(tx1)参照), 主効果Aが有意であり (F(2,98)=12.513, p=0, np2=0.203, 1-β=0.999), 主効果Bが有意でなく (F(2,98)=0.13, p=0.878, np2=0.003, 1-β=0.079), 交互作用AxBが有意であった (F(4,196)=13.941, p=0, np2=0.222, 1-β=0.999)。
主効果Aの検出力 (1-β) は十分である。交互作用AxBの検出力も十分である。なお検出力の値は水準間の相関係数に正負が混在している場合は平均相関を0と仮定し, それ以外はFisherの重み付きz変換値による平均相関を用いて算出した。
有意性を示した自由度2以上の効果についてMauchlyの球面性検定を行った結果 (Table(tx9)参照), 主効果Aについては有意でなかったことを確認した (Mauchly's w=0.993, p=0.845)。交互作用AxBについても有意でなかったことを確認した (Mauchly's w=0.885, p=0.76)。

⇒js-STARの第二枠『交互作用AxBが有意のときに』を実行してください。

> # オプション: [t] ⇒行頭の#を消す⇒ [Enter]
> # windows();boxplot(data~A*B,SAB) # 箱ひげ図
> # txD # Median,歪度,尖度,SEなど
> # TesA(水準=1, 相手=2) # 主効果Aの2水準 t検定
> # TesB(水準=1, 相手=2) # 主効果Bの2水準 t検定
> # write(txt,file.choose(),ap=T) # 結果のファイル保存

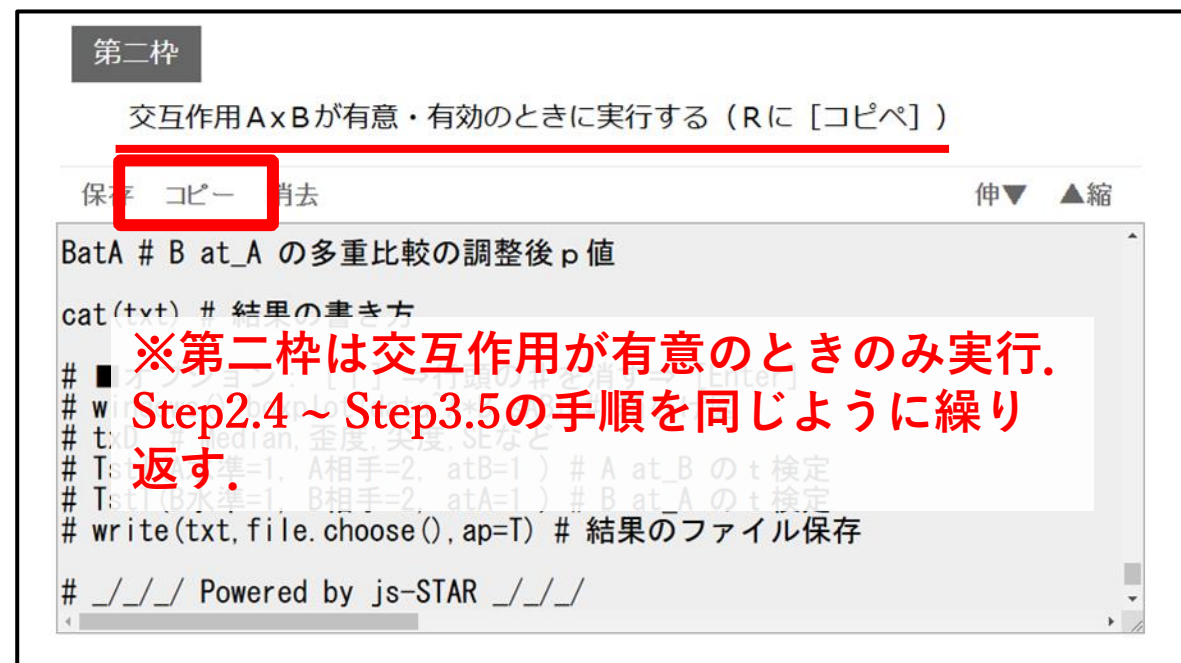
> # _/_/_/ Powered by js-STAR _/_/_/
```

結果の書き方が出力されるので, これを参考にして論文などに記述可能.

※図は別途, 自作する必要あり

図4.10 : RStudioで実行後のConsoleへの出力

Step3.6 : Rプログラムの第二枠のコピー



```
第二枠

交互作用AxBが有意・有効のときに実行する (Rに [コピー])

保存 コピー 消去 伸▼ ▲縮

BatA # B at_A の多重比較の調整後p値

cat (+v+) # 結果の書き方

# ■
# w
# tx
# T
# T
# write(txt, file.choose(), ap=T) # 結果のファイル保存

# _/_/_/ Powered by js-STAR _/_/_/
```

※第二枠は交互作用が有意のときのみ実行.
Step2.4 ~ Step3.5の手順を同じように繰り返す.

図4.11 : js-STAR XR+でのRプログラムの第二枠のコピーと実行に関する注意