

【2-20】 分散分析と多重比較

Meraki&生物基礎【SSH】 測定データの統計解析

研究を深めると条件の増加が起きます。3つ以上の条件で有意差検定を行う方法を紹介します。

統計解析（上級生編）

t検定（条件2つ）

分散分析・多重比較（条件3つ以上）

t検定が終わったらチャレンジ！

t検定の数式は
条件2つについて平均や分散を用いたね
条件が3つ以上では
他の方法を使います。なぜだろう？



0 t検定は条件2つまで！！ なぜ！？

11/17

強豪チームA ↔ 弱小チームC

対戦成績 Aの19勝1敗 勝率.950

11/24

強豪チームB ↔ 弱小チームC

対戦成績 Bの19勝1敗 勝率.950

2試合で A○-C●、B○-C●になる確率（推定）は

$$0.95 \times 0.95 = 0.9025$$

勝率 95%のゲームが
2試合続いたとしても、
勝率の高いチームが
2試合とも勝利する確率は
95%ではないよね



	平均	標準偏差	
15°C	2.66	0.15	P<0.05
20°C	2.77	0.33	
25°C	4.00	0.25	

t検定で、15°Cと25°Cで95%水準の有意差
t検定で、20°Cと25°Cで95%水準の有意差

15°C・20°Cに対して25°Cの値が95%水準で
有意に高いとはいえない。

左の結果では
理論的には「25°Cが15°Cと
20°Cに対して有意に高くな
る水準は 95%ではなく、
90.25%」になるんだ

t検定を行えば行うほど
水準は下がっていくよ



各温度におけるばらつき

	測定値
15°C	2.0 2.1 1.9 2.0 1.8 1.7
20°C	2.4 2.1 2.2 2.0 1.5 1.8
25°C	3.9 3.7 4.1 4.3 3.7 3.9

温度をまた
いで
の
ばらつき

この値が大きければ条件
間の有意差に近づく。

t検定は「条件2つずつに絞
って分析する」のが特徴です

条件3つ以上の結果を分析
するなら、「条件3つ以上の
データをまとめて分析する」
方法が必要になります



分散分析シート ☆ ☆ ☆

ファイル 編集 表示 挿入 表示形式 データ ツール 拡張機能 ヘルプ

100% 123 MS ...

分散分析シート

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	分散分析シート									←網掛けのセルには自分たちの値を入力しよう。			
2	①元データを入力					②各条件の合計				③④⑤の各条件の値を各データ①の(条件数から1を引いた)で割る			
3	測定値(複数)					合計				その2乗			
4	案件1	12.2	18.8	18.2		49.2	2420.64						
5	案件2	22.2	20.5	14.6		57.3	3283.29						
6	案件3	20.8	19.5	26.3		66.6	4436.56						
7	案件4	26.4	32.6	31.3		90.3	8154.09						
8	案件5	24.5	21.2	22.4		68.1	4637.61						
9						331.5							
10	各条件のデータ数は					3							
11	条件数は					5							
12	全データ数は					15							
13	④(①を2乗)					⑤(④の合計)				③④⑤の合計 - (②の合計) ÷ 条件数 = (全データ数 - 1) × ①(③) ÷ ⑤			
14	測定値(複数)の2乗					データの2乗和				変動要因 平方和 自由度 平均平方			
15	案件1	148.84	353.44	331.24	0	833.52							
16	案件2	492.84	420.25	213.16	0	1126.25							
17	案件3	432.64	380.25	691.69	0	1504.58							
18	案件4	696.96	1062.76	979.69	0	2739.41							
19	案件5	600.25	449.44	501.76	0	1551.45							
20					合計	7755.21							

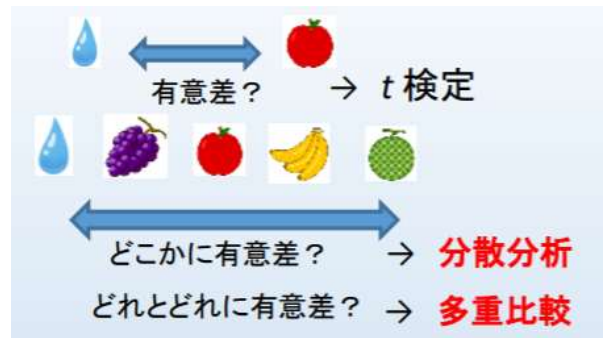
水瀬 7.122 一値が大きいほど(水瀬)

2 多重比較

分散分析は「複数の条件のどこかに有意差がある」

多重比較は「複数の条件のうち、具体的にどの条件とどの条件の間で有意差がある」

分散分析で有意差が得られた場合、多重比較を行います。



(1)帰無仮説 : 複数の条件で値に差はない。(条件3つであれば、例えば次の式が成り立つ。)

$$\frac{(\text{条件1の平均値}) + (\text{条件2の平均値})}{2} - \frac{(\text{条件3の平均値})}{1} = 0$$

例3 4月生まれ、10月生まれ、3月生まれの運動直後の心拍数を調べた。「3月生まれは循環器の発達が遅く、他の誕生日より心拍数が高くなる」と考えた。

	運動直後の心拍数 (回/30秒)				平均	係数
4月	71	68	74	68	70.25	1/2
10月	71	69	67	72	69.75	1/2
3月	67	72	70	73	70.5	-1/1

係数は、自分の仮説に基づいて全体が0になるように正・負の分数を割り当てます
もし4つの条件を考えて、1つの条件が高いと予想すれば、係数は「1/1, -1/3, -1/3, -1/3」です

帰無仮説の数式に入れると、

$$\frac{70.25 (\text{4月平均}) + 69.75 (\text{10月平均})}{2} - \frac{70.5 (\text{3月平均})}{1} = -0.5 (\text{絶対値は0.5})$$

となります。値が0に近く、有意差は得られないかもしれません。



(2)対立仮説 : 複数の条件で値に差がある。

例 運動直後は心拍数が高くなり、運動前や数分後に比べて心拍数が高くなると考えた場合、

	運動直後の心拍数 (回/30秒)				平均	係数
運動前	?	?	?	?	??	1/2
運動直後	?	?	?	?	??	-1/1
5分後	?	?	?	?	??	1/2

自分たちの測定データを使ってチャレンジしてみよう!

計算式の値は大きくなり、有意水準を超える値を得られるかもしれません。



※活動2

多重比較シート (元ファイルの別シート)

を用いて作業をしてください、

※分散分析と同じデータを使い、

(運動前、運動直後、5分後)

仮説に基づいて係数を考えてください。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O				
1	多重比較シート																		
2	←網掛けのセルには自分たちで値を入力しましょう。																		
3	①元データを入力							②平均を確認		③係数を設定 (帰無仮説による)									
4	測定値 (複数)				合計	データの2乗		平均	係数	②×③	③のデータ数				分分散				
5	条件1	12.2	18.8	18.2	49.2	2420.64	16.4	0.333	5.467	0.0370					平方和				
6	条件2	22.2	20.5	14.6	57.3	3283.29	19.1	0.333	6.367	0.0370	変動要因				平方和				
7	条件3	20.8	19.5	26.3	66.6	4435.56	22.2	0.333	7.400	0.0370	水準間変動				平方和				
8	条件4	26.4	32.6	31.3	90.3	8154.09	30.1	-0.500	-15.050	0.0833									
9	条件5	24.5	21.2	22.4	68.1	4637.81	22.7	-0.500	-11.350	0.0833									
10								合計	0.000	-7.167	0.2778					分分散			
11	条件1つのデータ数は							3	↑「=-1/2」「=1/3」と入力します。							変動要因	平方和		
12	条件数は							5								水準間変動	平方和		
13	全データ数は							15								水準間変動	平方和		
14	④ (1)を2乗							⑤ (2)の合計											
15	測定値 (複数) の2乗							データの2乗和											
16	条件1	148.84	353.44	331.24	0	833.52									▼f (F分布の値を調べる)				
17	条件2	492.84	420.25	213.16	0	1126.25									m/2		4		
18	条件3	432.64	380.25	691.69	0	1504.58									F(m/2,n/2)		3.478		
19	条件4	696.96	1062.76	979.69	0	2739.41									F値×数式=検定に用いる値				
20	条件5	600.25	449.44	501.76	0	1551.45									F値×数式=		6.564		
21								合計	7755.21									今回の値は...	
22																			
23																			
24																			

今回のは... 7.167 > 1

3 取り組んだことのチェック

活動3 Google の回答フォームに回答して、どこまで達成できたかを確認しましょう。

ループリック評価

項目 No.	チェック	内容
1	☐	t 検定（運動前後）を行い、有意差を確認することができた。
2	☐	3 つ以上の条件間では t 検定を使わない理由を理解できた。
3	☐	分散分析の F ₀ 値の求め方を理解できた。
4	☐	分散分析シートで値を求め、条件間に有意差があることを確かめた。
5	☐	分散分析と多重比較による検定の違いを理解できた。
6	☐	多重比較の帰無仮説で係数の与え方を理解できた。
7	☐	多重比較シートで分析し、条件のどれとどれに有意差があるかを確かめた。

4：非常に良い	3：しっかり取り組んだ	2：よく取り組んだ	1：取り組んだ
No.7 までを確かめた上で、独自に条件を考えて No.1,4,7 のいずれかの検定を行った。	No.7 までを確かめることができた。	No.4 までを確かめることができた。	No.1 までを確かめることができた。

4 グラフの作り方

①多重比較シートで用いたいデータに注目

	測定値				平均	係数
運動前	37	38	40	-0.5	37.75	-0.5
運動直後	71	68	77	1	72.75	1
5 分後	32	34	31	-0.5	31.5	-0.5

今回の値は…		
38.125	>	5.768
多重比較の値（絶対値）		検定に用いる値

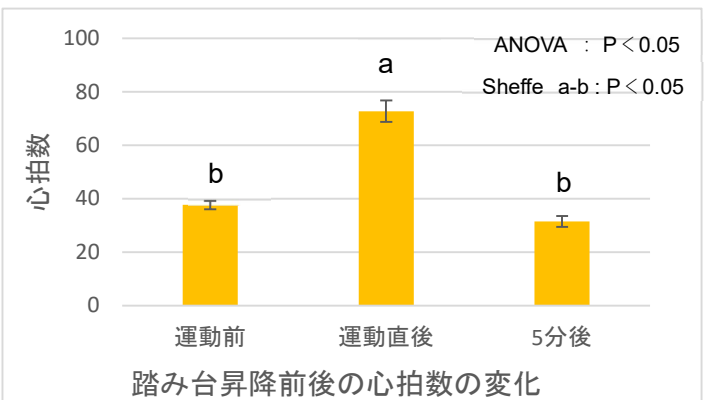
②標準偏差を作成（=stdev、グラフに使用）

	平均	標準偏差
運動前	37.75	1.708
運動直後	72.75	4.031
5分後	31.5	2.082

④有意差を a、b の記号で表現（有意差があるものは別記号、ないものは同じ記号）検定の方法、水準を記載

分散分析「analysis of variance」→「ANOVA」
多重比較「シェフェの方法を使った」→「Sheffe」

③作成したグラフを Word or PowerPoint に貼り付け

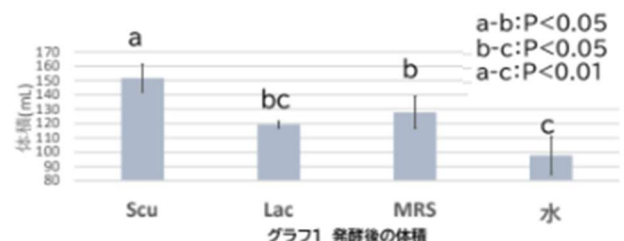


※多重比較シートの係数を変えながら様々な組合せで検定できますが、

解析ソフトを用いると、より複雑な比較も可能です。

68 期生の実験「ラクトースではパンが膨らまない」ことを検証したときの結果です。

手法	水準1	水準2	平均1	平均2	P 値	*: P<0.05 **: P<0.01
Scheffe	Scu	Lac	151.7	119.3	0.0271 *	
	Scu	MRS	151.7	127.7	0.1034	
	Scu	水	151.7	97.8	0.0013 **	
	Lac	MRS	119.3	127.7	0.7871	
	Lac	水	119.3	97.8	0.1508	
	MRS	水	127.7	97.8	0.0394 *	



一般的な糖（スクロース）で有意に体積が大きく、ラクトースは MRS（発酵以外の栄養成分）、対象実験（水）の中間値、どちらに対しても有意差がない。ことを示しています。（MRS と水には有意差が現れています）