第5章　分散分析

教育学研究科　修士1年
山口将典

予備知識：要因・水準

質的な独立変数は要因（factor），

質的な独立変数の値は，その要因の水準（level）とよばれる。

（南風原，2002, p.264）

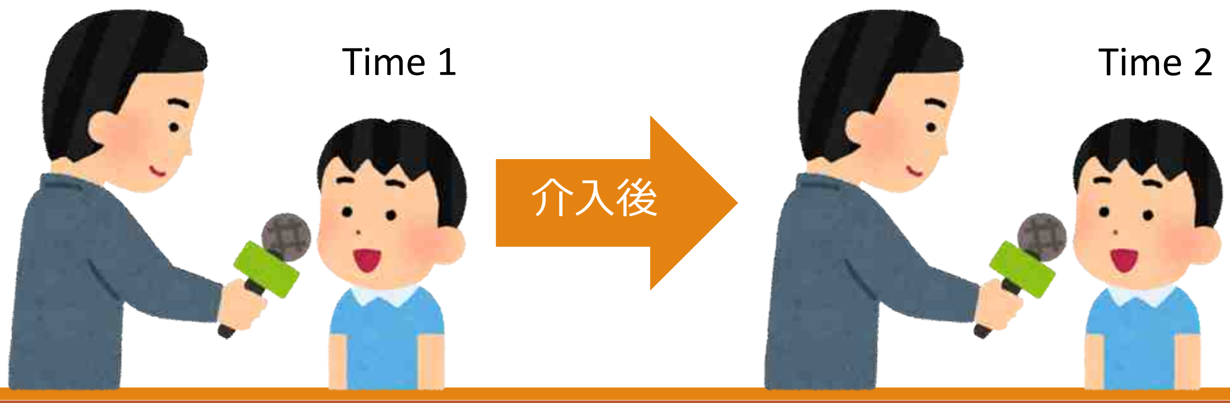
例えば，「性別」という要因は，「男性」および「女性」という水準をもつ。

予備知識：対応のある/ない要因

要因の中でも「ある条件におけるある被験者の測定値の高低によって，その被験者の別の条件における測定値の高低がある程度予想できる」ものを「対応のある要因」とよぶ。

「対応のない要因の場合は，その要因の異なる水準に含まれる従属変数の値は互いに独立であるのに対し，対応のある要因の場合は，異なる水準に含まれる従属変数の値に相関がある」

（南風原, 2002, p.267）



（e.g. 測定時期という要因）

予備知識：被験者間/内要因

被験者内要因：各水準に同一の被験者が割り当てられる

被験者間要因：各水準に異なる被験者が割り当てられる

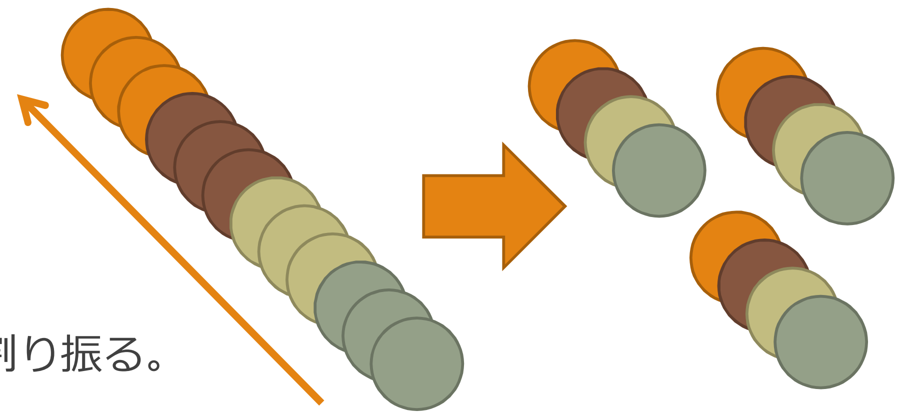
※被験者間要因であっても，対応のある要因になることがある。

例えば，ある得点に基づいて実験の

条件数と同じ人数からなる得点の似通った

被験者の対（ブロック）を作り（マッチング），

ブロックごとに被験者を異なる条件にランダムに割り振る。



予備知識：重要なのは，対応のある/なし

・対応のある要因の方がその要因の水準間の比較可能性が高まる。

→ある水準のみに特定の傾向を持つ人が集中することを避けることが出来る

・対応のある要因のほうが，一般的にその効果が検出できる確率（検定力）が高まる

→対応づけたほうが，母集団に存在する平均値差が，正確に標本平均値差に反映される可能性が高まる。

（南風原, 2002）

・ただし，被験者の負担等の観点から被験者内要因か被験者間要因かも重要。

マッチング：なぜ検定力が高まるのか

対応がある要因（変数間に正の相関がある）の場合，共分散は正の値をとる($r = \frac{s_{xy}}{s_x s_y}$)

→相関が高いほど，共分散は大きくなる($r = \frac{s_{xy}}{s_x s_y}$)。

2つの変数xとyの差の分散（${\sigma^2}_{x-y}$）を考えると，

$${\sigma^2}_{x-y} = {\sigma^2}_x - 2\sigma_{xy} + {\sigma^2}_y$$

このとき，標準偏差は，

$$\sigma_{x-y} = \sqrt{{\sigma^2}_x - 2\sigma_{xy} + {\sigma^2}_y}$$

したがって，共分散が大きい（正の相関がある，対応のある要因）ほど，

標準誤差（$\frac{\sigma_{x-y}}{\sqrt{N}}$）は小さくなるので，検定力が高まる。

分散分析：定義

「質的な独立変数の値によって従属変数の**平均**がどのように異なるかを分析するための方法」

（南風原，2002, p.263）

「群の**平均値**の比較を，全体のバラツキ（分散）を2つのバラツキ（分散）に分解して分析しようとするのが，分散分析です」

（山田・村井, 2004, p.163）

帰無仮説　H_0: $\mu_1 = \mu_2 = \cdots = \mu_a$

対立仮説　H_1: 少なくとも1つは不等号

分散分析とは：なぜ分散で平均値の比較をするのか

分散とは，データのバラつきを表す指標（濃くなっている部分を全体平方和と呼ぶ）

$$s^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}$$

「分散分析という名称は，『データのばらつき（**分散**）を，群間の違いという意味のはっきりしている平方和（群間平方和）と誤差平方和に分解して（**分**けて）解**析**する』と理解しておくのがよい」

（永田・吉田, 1997, p.7）

分散分析とは：平方和の分解

	測定値のバラつき	=	実験要因の効果によるバラつき	+	誤差によるバラつき
被験者間計画 1要因	全体平方和 (データの値－全平均)^2	=	群間平方和 （群平均－全平均）^{2}	+	群内平方和 (データの値－群平均)^2
被験者内計画 1要因	全体平方和 (データの値－全平均)^2	=	条件平方和 （条件平均－全平均）^{2}	+	個人差による平方和 （各被験者の平均－全平均）^{2} ＋誤差平方和

分散分析：利点①

一発で「どこかに平均値差があるか」がわかる。

→水準が多い実験で，探索的に群間差を調べる場合，群のペアをそれぞれ検定するのは大変

例えば，水準が5つある場合，$_5C_2 = 10$通りもやらなくてはならない…

というよりも，

そもそも2群間の差の検定を何度も行ってはいけない！

2群間の平均値差の検定の繰り返しはダメ？

「すべての水準対で5%水準で2群間の平均値差の検定をする方法だと，全体としての帰無仮説が正しいとき，対に関する帰無仮説が少なくとも1つ棄却され，結果的に全体としての帰無仮説が棄却される確率は5%よりも大きくなってしまう」

（南風原, 2002, p.281）

H_0: $\underline{\mu_1 = \mu_2} = \mu_3 = \cdots = \mu_a$

赤くなっている部分に差が合った場合，論理的帰結として帰無仮説全体が棄却される

例えば，水準が5つのとき必要な検定は10回。10回とも有意にならない確率は0.95^{10}。

少なくとも1回有意になる確率は$1 - 0.95^{10} = 0.4013$(約40%！)。

分散分析：利点②

要因が2つ以上のとき，単一の要因の効果（主効果）以上の効果（交互作用効果）を見いだせる。

主効果　　　　　　　：従属変数に影響を及ぼしている単一の要因の効果

交互作用（効果）：従属変数に対して，本来互いに独立の２つ（以上）の変数が，互いに関連し合って影響を及ぼしている複合的要因の効果

分散分析：交互作用を調べる

交互作用があるかどうかを簡易的に知りたいのであれば，「折れ線グラフを描いてみて，複数の線分が平行移動して重なり合うかどうかを調べてみればよい」（テキスト，p.79）。

→分散分析表における交互作用の項目は，折れ線が平行線から有意に隔たっているかどうかを確かめている。

単純主効果の検定

分散分析の結果，交互作用が有意であった場合，ある要因の各水準における別の要因の効果（単純主効果）を調べることになる。

→単純主効果の検定

「単純主効果の検定はすべての場合について行う必要はなく，研究の目的や仮説に応じて必要なものだけを行えばよい。なお，単純主効果の検定は，本来，交互作用が有意である場合のみ行われるものである」

（森・吉田, 1990）

⇔計画比較（もともと特定の群比較に興味があり，分散分析で有意か否かに関わらず，単純主効果の検定をする）という方法がないこともない…

多重比較

分散分析および単純主効果の検定でいずれかの要因で有意差が出た（どこかに差があった）場合，

次に知りたいのは「どこに差があったのか」。

→多重比較法で調べる。

多重比較法とは，「検定を複数回繰り返すことに対して調整を施した手法」である。

（永田・吉田, 1997, p.i）

先述のように，t検定を繰り返し行うことは帰無仮説が棄却される確率を上げてしまう。そこで求めた*p*値に対して何らかの調整を行う。

調整方法にはいくつかの種類があり，データの種類や検定の目的によって選択する。

多重比較法の選択
（永田・吉田, 1997, p.8）

対比較：

2群間での比較。

単調性：

各群に順序付けがあり，母平均が単調増加あるいは単調減少していると想定する。単調変化する群平均のなかのどこで有意な差が生じているかを調べたい場合などに用いられる。

対比：

以下のような帰無仮説を考える

$$H_0\colon \frac{\mu_1+\mu_2}{2}=\mu_3$$

$$H_0\colon \frac{\mu_1+\mu_2}{2}=\frac{\mu_3+\mu_4}{2}$$

表1.3　本書で解説する多重比較法の適用の場と手法の特徴

手法名（節）	適用状況					手法の特徴			
	すべての群間の対比較	対照群との対比較	単調性を想定した対比較	対比についての検定	その他	パラメトリック法	ノンパラメトリック法	シングルステップ法	ステップダウン法
テューキー（4.2節）	□					○		○	
ダネット（4.3節）		□				○		○	
ウィリアムズ（4.4節）			◎			○			○
シェフェ（4.5節）	△	△	△	○		○	○	○	
スティール・ドゥワス（5.2節）	□						○	○	
スティール（5.3節）		□					○	○	
シャーリー・ウィリアムズ（5.4節）			◎				○		○
ボンフェローニ（6.1節）	△	△	△	△	△	○	○	○	
ホルム（6.2節）	□	□	□	○	○	○	○		○
シェイファー（6.3節）	○			(○)	(○)	○	○		○
ホランド・コペンハーバー（6.3節）	○			(○)		○	○		○
テューキー・ウェルシュ（7.2節）	○					○	○		○
ペリ（7.3節）	◎					○	○		○
ダネットの逐次棄却型検定（7.5節）		◎				○	○		○

検出力の良い順に

◎＞○＞□＞△

多重比較法の選択
（永田・吉田, 1997, p.8）

パラメトリック：

２つの仮定を置く。

仮定１：各群の母集団分布が正規分布である

仮定２：すべての群を通して母分散は等しい（等分散性）

ノンパラメトリック：

必ずしも正規分布を想定しない。

ただし，a個の群の母集団について考える際，各分布の形状は同じで，その位置のみが異なっている場合を想定する。

→等分散性は仮定していることに注意！

→サンプルサイズが小さいからノンパラメトリック法というのは正しくない。

検出力の良い順に
◎＞○＞□＞△

表1.3　本書で解説する多重比較法の適用の場と手法の特徴

手法名（節）	適用状況					手法の特徴			
	すべての群間の対比較	対照群との対比較	単調性を想定した対比較	対比についての検定	その他	パラメトリック法	ノンパラメトリック法	シングルステップ法	ステップダウン法
テューキー（4.2節）	□					○		○	
ダネット（4.3節）		□				○		○	
ウィリアムズ（4.4節）			◎			○			○
シェフェ（4.5節）	△	△	△	○		○	○	○	
スティール・ドゥワス（5.2節）	□						○	○	
スティール（5.3節）		□					○	○	
シャーリー・ウィリアムズ（5.4節）			◎				○		○
ボンフェローニ（6.1節）	△	△	△	△	△	○	○	○	
ホルム（6.2節）	□	□	□	○	○	○	○		○
シェイファー（6.3節）	○			(○)	(○)	○	○		○
ホランド・コペンハーバー（6.3節）	○			(○)		○	○		○
テューキー・ウェルシュ（7.2節）	○					○	○		○
ペリ（7.3節）	◎					○	○		○
ダネットの逐次棄却型検定（7.5節）		◎				○	○		○

効果量：帰無仮説検定の問題

「帰無仮説は常に間違っている」（大久保・岡田, 2012, p.32）

実際の関心は「要因の効果があるかどうか」よりも「要因の効果の大きさ」であることが多い。

効果量はそのような要因の効果の大きさを示す指標。

分散分析で用いられるη^2は，分散説明率（あるいは決定係数; 従属変数の分散のうち説明変数によって何パーセントが説明できるか）を表す。

効果量：η^2と$\eta_p{}^2$の違い

$$\eta^2 = \frac{\text{要因}_A\text{効果の平方和}}{\text{平方和の総和}}$$（厳密には標本効果量なので$\hat{\eta}^2$）

→η^2は分散説明率（測定値の分散を要因の分散でどれくらい説明できるか）を表す。

$$\eta_\mathrm{p}{}^2 = \frac{\text{要因}_A\text{効果の平方和}}{\text{要因}_A\text{効果の平方和}+\text{誤差平方和}}$$（厳密には標本効果量なので$\hat{\eta}_\mathrm{p}{}^2$）

→ $\eta_\mathrm{p}{}^2$の分母は全分散から関心のない効果を除去した量。

効果量：η^2とη_p^2の違い

η^2は要因が多いほど分母が大きくなるので，要因が多いほど小さくなる。

→研究間で比較するには不都合（ただし１要因被験者間分散分析は例外）。

η_p^2でも被験者間のデザインで行われた研究と，被験者内のデザインで行われた研究とを比較することはできない。

→「被験者内のデザインでは分母から被験者効果の分散を除くことが出来るため，被験者間の場合と比べてη_p^2が大きな値をとる」（大久保・岡田, 2012, p.89）

→一般化効果量というものが提唱されていたり…

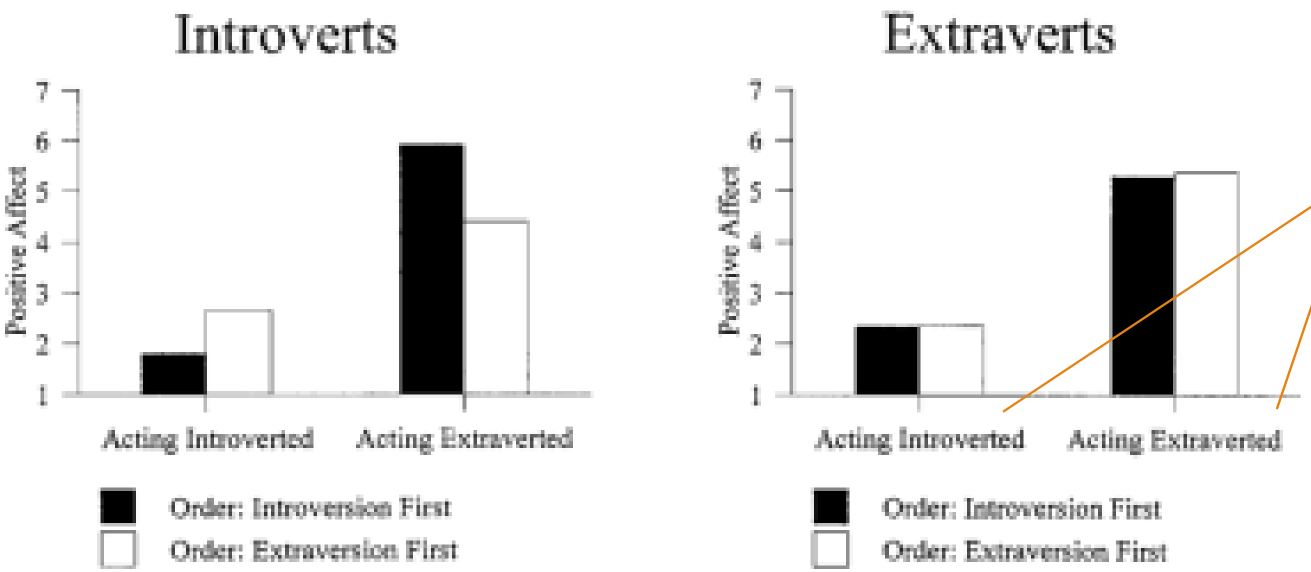
（大久保・岡田, 2002, p.89-92参照）。

Zelenski, Santoro, Whelan, 2012の研究

背景：外向的に振る舞ったほうがポジティブ感情が高くなる

Fleeson, Malanos,& Achille, 2002の研究より（Figure 5）

要因：気質（外向性）
水準：内向的or外向的
被験者間要因（対応なし）



要因：教示
水準：外向的に振る舞う
　　　内向的に振る舞う
被験者内要因（対応あり）

Zelenski et al.(2012)の問い：

自分の気質と異なる振る舞いをするのって疲れるのでは？
→自我消耗（ego depletion）の度合いで確かめてみよう

Fleeson et al.（2002）では教示は被験者内計画だった。
→教示も被験者間計画でやったほうが影響がない。

方法：

参加者：

心理学専攻の大学生127名

気質評価テスト（オンラインで実施）→上位30%＝外向気質（n=60）

→下位30%＝内向気質（n=67）

インタビュー：参加者のことを知るという名目（休日は何をしているか etc.）。インタビューを受ける参加者には事前に，外向的，内向的，あるいは自然に振る舞うよう教示を与えた。インタビュアーは誰がどの教示を受けたか知らない。

方法：

ストループ課題：

赤　青　黄　緑

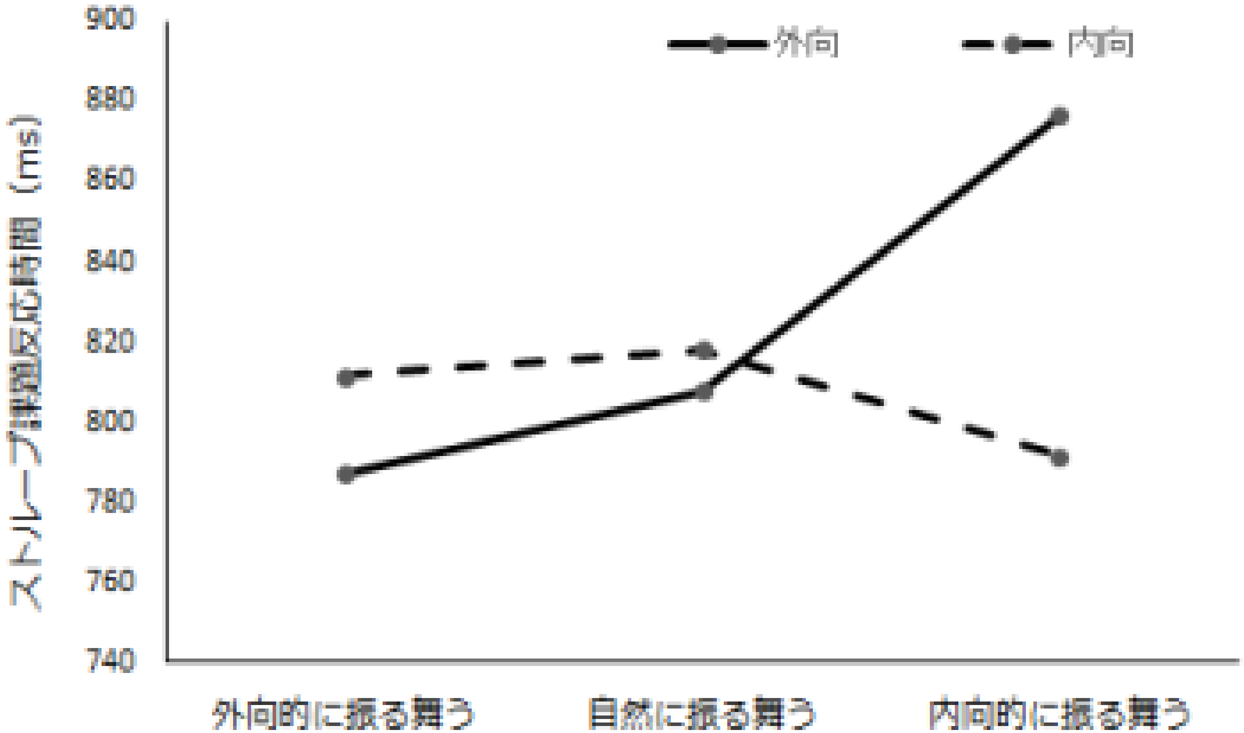
不一致試行（249試行）

一致試行　（25試行）

→正解したときの反応時間を計測

反応時間が300ms以下と3000ms以上である場合を除いて，その平均値をだす。

結果：グラフにしてみた…これって意味のある差？



実際に使ってみる

RとANOVA君を使って

Rの起動（ANOVA君搭載）

・MLで送ったテキストファイル（anovakun_480）,csvデータ（９章__サンプルデータ1）を保存する。

・左下のWindowsのアイコン を押してRを検索し，起動する。

・「ファイル」→「Rソースの読み込み」を選択。

・右下の「R files *.R」を「All files(*.*)」に変更し，保存したテキストファイル（anovakun_480）を開く。

・「> source(“C://Users//…//anovakun_480.txt”)」と表示される。

csvファイルからデータをコピー

・先程保存したcsvファイル（９章__サンプルデータ1）をexcelで開く。

・excelのB2からD61までをコピーする。

Rでデータを読み込む

・一番下の行が「>」となっているので，それに続けて，

「>　dat<-read.table(“clipboard”)」と入力する。

→これでRにデータが読み込まれたはず（> datと入力すると，読み込んだデータを確認できる）。

分散分析をする

コードは，次のように入力する。

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

「データ」…分析したいデータを参照する変数の名前です。データを読み込むときにつけた変数名を入力する（datなど）。

「要因計画」…「“（被験者間要因のラベル）s（被験者内要因のラベル）”」というように入力する（ラベルはデータ列の左側からA,B,C,…Zとなる）。

「各要因の水準数」…要因計画で入れた順番通りに水準数を入力する。水準数の間はカンマ（,）で区切る。

・今回の場合は「> anovakun(dat,”ABs”,2,3)」と打ち込み，エンターを押す。

結果の表示

<< DESCRIPTIVE STATISTICS >>

→記述統計（要因水準ごとに被験者数，平均値，SD）が確認できる。

<< ANOVA TABLE >>

→分散分析表が確認できる。

SSは平方和，dfは自由度，MSは平均平方和，F-ratioはF比（値），p-valueはp値。

結果：分散分析表の見方（テキストp.82）

ソース	平方	η_p^2
気質	12988.	0.0130
教示	37765.	0.0369
気質×教示	125087	0.1127
誤差	984775	
総和	116061	

（注） ns=non-sig

ソースとは，分散の発生源

「気質」「教示」：従属変数に対して影響を及ぼしている単一の要因
→主効果

「気質×教示」：従属変数に対して，本来互いに独立の２つ（以上）の変数が，互いに関連し合って影響を及ぼしている複合的要因
→交互作用

「誤差」：実験要因によらない影響

結果：分散分析表の見方（テキストp.82）

ソース	平方和	自					η_p^2
気質	12988.3421					112	0.0130
教示	37765.9777	2	18882.9890	1.0354	0.362ns	0.0325	0.0369
気質×教示	125087.2606					078	0.1127
誤差	984775.6019						
総和	1160617.1823						

（注） ns=non-significant, †: $p<.10$, *: $p<.05$, **: $p<.01$, ***: $p<.001$

単一要因の平方和＝（群平均－全平均）²の合計

２つの要因を仮にj，k，測定値をyとすると，

$$\text{交互作用の平方和} = n\sum_{j=1}^{a}\sum_{k=1}^{b}(\bar{y}_{jk} - \bar{y}_{j\cdot} - \bar{y}_{\cdot k} + \bar{y})^2$$

誤差の平方和＝（データの値ー群平均）²の合計

結果：分散分析表の見方（テキストp.82）

ソー[illegible]	[illegible]	自由度	平均平方	F値	p値	η^2	η_p^2
気質	[illegible]	1	12988.3421	0.7122	0.4024ns	0.0112	0.0130
教示	37765.9777	2	18[illegible]	[illegible]	[illegible]	[illegible]	0.0369
気質×教示	125087.2606	2	62543.6303	3.4296	0.0396*	0.1078	0.1127
誤差	984775.6019	54	18[illegible]				
総和	1160617.1823	59					

（注）[illegible]t, †: $p<.10$, *: $p<.05$, **: $p<.01$, ***: $p<.001$

単一要因の自由度
＝水準の数－1

交互作用の自由度
＝(要因Aの水準数－1)(要因Bの水準数－1)

誤差の自由度
＝参加者数－(要因Aの水準数)(要因Bの水準数)

総和の自由度
＝参加者数－1

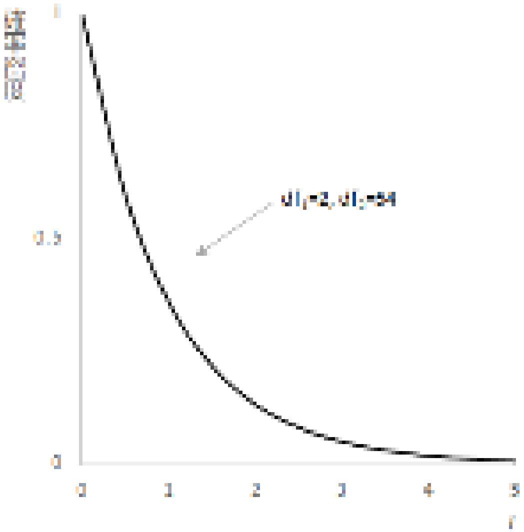
結果：分散分析表の見方（テキストp.82）

ソース	平方和		自由度		平均平方	F値	p値	η^2	η_p^2
気質	12988.3421	÷	1	＝	12988.3421	0.71			
教示	37765.9777	÷	2	＝	18882.9890	1.03			
気質×教示	125087.2606	÷	2	＝	62543.6303	3.4296	0.0396*	0.1078	0.1127
誤差	984775.6019	÷	54	＝	18236.5852				
総和	1160617.1823		59						

（注） ns=non-significant, †: $p<.10$, *: $p<.05$, **: $p<.01$, ***: $p<.001$

平均平方
＝各平方和を各自由度で割ったもの

結果：分散分析表の見方（テキストp.82）

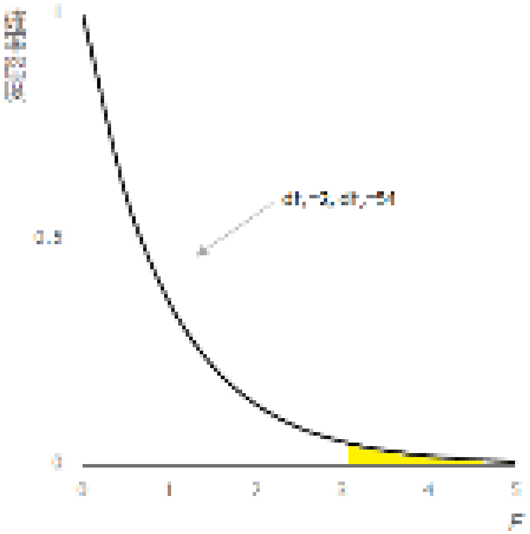


自由度	平均平方	*F* 値	*p* 値	η^2	η_p^2
1	12988.3421	**0.7122**	0.4024ns	0.0112	0.0130
2	18882.9890	**1.0354**			
2	62543.6303	**3.4296**			
54	18236.5852				
59					

（注）　ns=non-significant, †: $p<.10$, *: $p<.05$, **: $p<.01$, ***: $p<.001$

$$F値 = \frac{\text{要因によるバラツキ}}{\text{誤差によるバラツキ}}$$

結果：分散分析表の見方（テキストp.82）



自			p値	η^2	η_p^2
			0.4024ns	0.0112	0.0130
2	18882.9890	1.0354	0.362ns	0.0325	0.0369
2	62543.6303	3.4296	0.0396*	0.1078	0.1127
54					
59					

帰無仮説は
「独立変数の効果がゼロ」

自由度がそれぞれ2，54のとき，
その確率密度が5%になるFの値は約3.16

（注） ns=non-significant, †: p<.10, *: p<.05, **: p<.01, ***: p<.001

結果の表示

<< POST ANALYSES >>

事後検定をしたことを示している。

< SIMPLE EFFECTS for “A x B” INTERACTION >

AとBの交互作用があったので，単純主効果検定をしたことを示している。

結果：単純主効果検定

ソース	平						
気質 at 外向的	790						
気質 at 内向的	434						
気質 at 統制	155						
教示 at 外向群	150138.6165	2	75069.3083	**4.1164**	**.0217***	.1294	.1323
教示 at 内向群	12714.6218	2	6357.3109	**0.3486**	**.7073ns**	.0110	.0127
誤差	984775.6019	54	18236.5852				

（注） ns=non-significant, †: p <.10, *: p <.05, **: p <.01, ***: p <.001

*F*値と*p*値をみると，
教示 at 外向群で有意な差があることがわかる
（$F(2,54)=4.1164$, $p=.0217$）
外向群では教示の仕方によって疲労度に違いがあった！

どの水準間に差があったのか？→多重比較

結果の表示

< MULTIPLE COMPARISON for "B at a1" >

a1（気質の外向群）についてB（教示）の各水準の群間差を多重比較法で検定していることを示している。

== Shaffer's Modified Sequentially Rejective Bonferroni Procedure ==

調整方法がShafferの方法であることを示している。

== The factor < B at a1 > is analysed as independent means. ==
ANOVA君では被験者間要因として分析されたことを示している。

== Alpha level is 0.05. ==

有意水準が5%であることを示している。

結果：多重比較

ペア	差分	t値	自由度	p値	調整済みp値	
外向的－内向的	-169.327	2.8038	54	0.007	0.021	外向的＜内向的*
内向的－統制	116.555	1.9299	54	0.5089	0.0589	内向的＝統制
外向的－統制	-52.772	0.8738	54	0.3861	0.3861	外向的＝統制

ここではシェイファーの方法が使われているらしい…

調整済みp値をみると,「外向的ー内向的」のみ有意確率5%水準で有意であることがわかる。

→外向群の参加者は,「外向的に振る舞うより内向的に振る舞った後の方がストループ課題での反応時間が長くなった」

分散分析結果に効果量を表示させる

「> anovakun(dat,“ABs”,2,3,peta=T)」

と入力すると，偏イータ2乗（η_p^2）が算出される。

「> anovakun(dat,“ABs”,2,3,eta=T)」

と入力すると，イータ２乗（η^2）が算出される。

その他にも色々できる。詳しくは井関龍太先生のANOVA君のページで。

http://riseki.php.xdomain.jp/index.php?ANOVA%E5%90%9B%2FANOVA%E5%90%9B%E3%81%AE%E4%BD%BF%E3%81%84%E6%96%B9

参考文献

テキスト：

荘島宏二郎(編) (2017). 計量パーソナリティ心理学. ナカニシヤ出版. pp.77-92.

その他：

Fleeson, W., Malanos,A. B., & Achille, N. M. (2002). An intraindividual process approach to the relationship between extraversion and positive affect: Is acting extraverted as “Good” as being extraverted? *Journal of Personality and Social Psychology*. *83* ,6, 1409-1422.

南風原朝和 (2002). 心理統計学の基礎―統合的理解のために. 有斐閣.

井関龍太のページ　ANOVA君の使い方

（http://riseki.php.xdomain.jp/index.php?ANOVA%E5%90%9B%2FANOVA%E5%90%9B%E3%81%AE%E4%BD%BF%E3%81%84%E6%96%B9）

（2017年4月20日確認）

森 敏昭・吉田寿夫 (1990). 心理学のためのデータ解析テクニカルブック. 北大路書房.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

大久保街亜・岡田謙介 (2012). 伝えるための心理統計―効果量・信頼区間・検定力―.　勁草書房

山田剛史・村井潤一郎 (2004). よくわかる心理統計. ミネルヴァ書房.

Zelenski. J. M., Santoro,M.S., & Whelan,D.C. (2012). Would introverts be better off if they acted more like extroverts? Exploring emotional and cognitive consequences of counterdispositional behavior. *Emotion*, *12*, 290-303.