入門 統計学　第10章

実験計画法

『入門　統計学　第2版　－検定から多変量解析・実験計画法・ベイズ統計学まで－』（オーム社）

※注：本書を購入された方へのサービスですので，教科書指定（参考図書は不可）していない授業での使用はお控えください。

実験計画法とは

- 良いデータを得るためには，どのように実験を組めば良いのかを考えること（実験計画法という分析手法があるわけではない）
- 分散分析などで正しい解析結果を得るためには，誤差の影響が少ない“良いデータ”が必要
- データを効率的に収集することも目的のひとつ

R.A. Fisher
(1890〜1962)

10.1　フィッシャーの三原則

実験を成功させるために従うべき3つの原則

どちらの過誤も犯さない

第一種の過誤：効果がないのにあると判断

第二種の過誤：あるべき効果を見逃す

あるべき要因効果を誤らずに検出

反　復　偶然誤差を評価し減らすため，群内で独立した実験を繰り返す

無作為化　系統誤差を減らすため，実験の場を“でたらめ”に配置する

局所管理　系統誤差と偶然誤差を減らすため，実験の場を小分けにし，その中で実験を管理する

10.2　原則その１：反復

（偶然誤差を減らす）

群（水準）内で独立実験を繰り返し，データを複数収集する

反復数を増やすと良いことがある？

変動計算に用いる群平均の偶然誤差が小さくなり，検定の精度が向上

群間変動が大きくなり，F値が大きくなって帰無仮説の棄却域に入りやすくなる

分散分析の検定統計量　F値＝$\dfrac{\text{群間変動}／（\text{群数}－１）}{\text{群内変動}／（\text{データ総数}－\text{群数}）}$

反復がなければ，そもそも分母の誤差を評価できない

自由度が大きくなり，限界値が小さくなるため検出しやすくなる

圃場における反復の事例

反復の原則が
守られていない圃場配置

反復の原則が
守られている圃場配置

収穫量に差があっても誤差の範囲内なのかどうか判別できない

理想的な反復数は検出力分析で求める（次掲）

（分散分析のための）反復数の計算

これが問題

反復数

効果量 f ：要因が持つ効果そのものの大きさ

有意水準α ：許容できる第一種の過誤を犯す確率

（α=0.05が一般的）

検出力（1-β） ：あるべき要因効果を検出できる能力

（0.8が一般的）

群　数 ：処理水準の数

群数も非心度に影響を与える

（第6章のt検定の検出力分析では不要だった）

分散分析の効果量と
ソフト（G*power)による反復数の計算

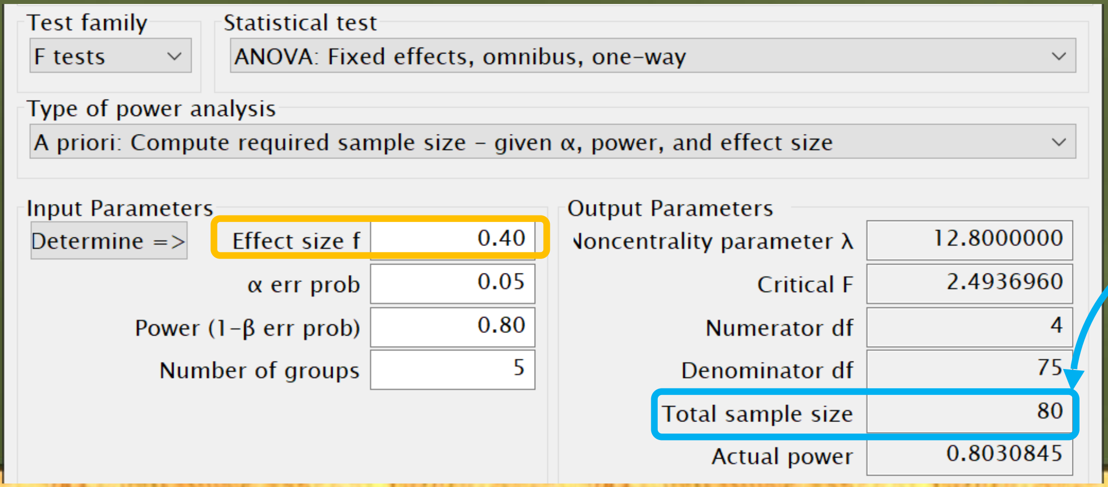
分散分析の効果量

$$\hat{f} = \sqrt{F\text{値} \times \boxed{\frac{\text{群数}}{\text{データ総}}}} = \sqrt{\frac{\text{群 間 変}}{\text{群内}}}$$

Fから自由度の影響を取り除いている

※手がかりがない場合：
大効果=0.4
中効果=0.25
小効果=0.1

5群の分散分析の反復数をG*powerで効果量0.4を想定して計算



←対応のない一元配置分散分析

5群合わせて80回：
独立した実験を1群あたり
16回反復させるべき

反復数の目安
（対応のない一元配置分散分析）

群（水準）数	効果量の大きさ（α=0.05, 検出力=0.80の場合）		
	小 (f=0.10)	中 (f=0.25)	大 (f=0.40)
2 群（t検定）	393	63	26
3 群	323	53	22
4 群	274	45	19
5 群	240	40	16
6 群	215	36	15
7 群	196	33	14

効果量が大きくなるほど反復数は少なくて良い

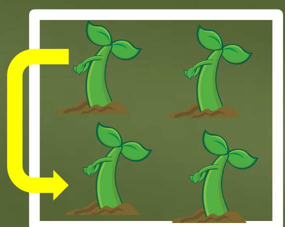
意外と沢山の反復が必要

群数が増えるほど反復数は少なくて良い

注意点（疑似反復）

- 反復とは独立した実験を繰り返し，異なる個体からデータを測定すること（replication）
- 処理を1度しか行わず，同じ個体からデータを繰り返し測定するのは疑似反復（repeat）：第一種の過誤を増大

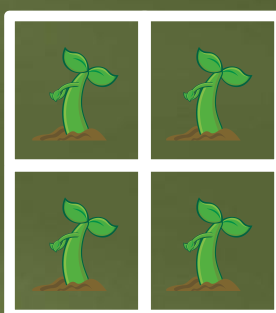
土が肥えていればどれも成長が良い（独立していない）



~~反復数＝4~~

反復数＝1

（4個体の平均）



反復数＝4

同じ土壌で栽培した苗を測定しても独立した実験とは言えない

コンクリートで仕切った区画で栽培すれば，土の肥え具合は独立する

10.3　原則その２：無作為化

（系統誤差対策）

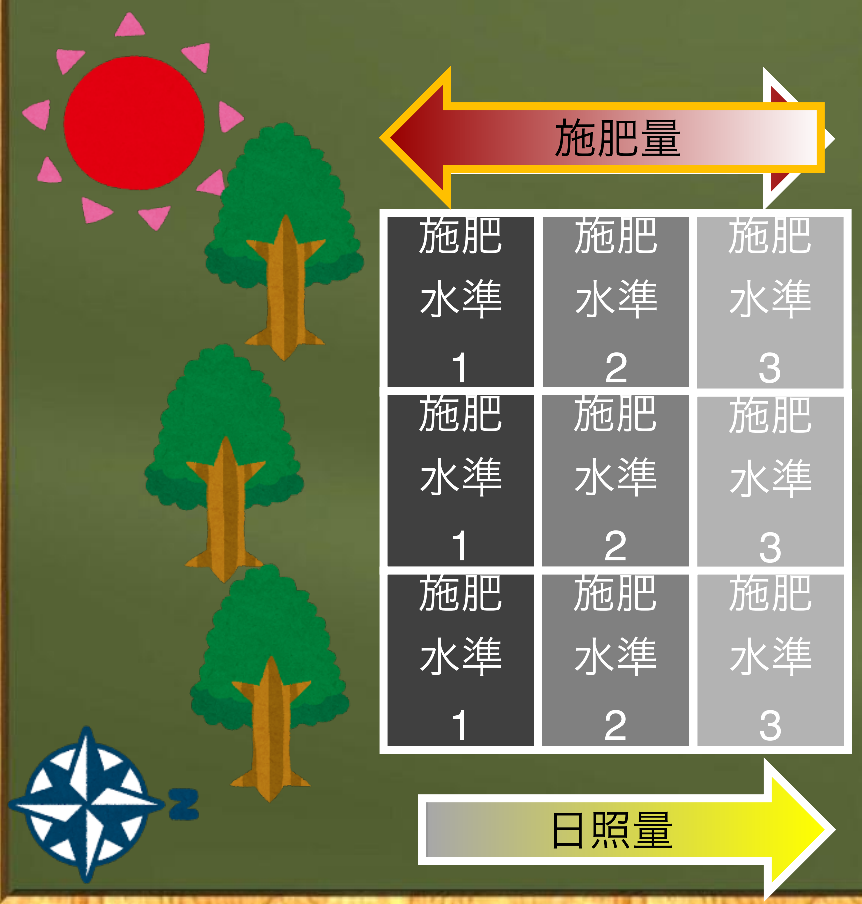
- 分散分析の結果を誤るときはどんなとき？

　→誤差効果が偏りを持って（系統誤差として）実験結果に入り込んでしまったとき

　→誤差効果が主効果と交絡して検定統計量を本来よりも大きく（小さく）歪める

- 対策：取り除くことは難しいので，実験の場の配置（区画・計測順）を無作為に並び替える

無作為化していない圃場実験例



（方向性を持っているため）分母にあるべき誤差が分子に来てしまう

検定統計量

(施肥量)　(日照量)

$$F = \frac{\boxed{\text{要因効果}} + \boxed{\text{系統誤差}}}{\boxed{\text{偶然誤差}}}$$

(施肥量以外の効果)

施肥効果と日照効果が交絡して検定統計量が大きい方へ歪む（第一種の過誤の増大）

施肥量と日照量が逆方向ならば小さく歪み，第二種の過誤が増大する

無作為化の導入

検定統計量

(施肥量)　　(日照量)

$$F = \frac{\boxed{\text{要因効果}} + \boxed{\text{系統誤差}}}{\boxed{\text{偶然誤差}}}$$

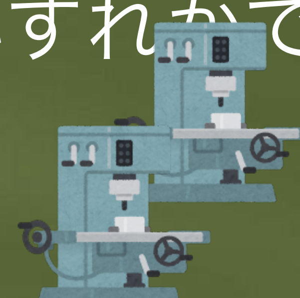
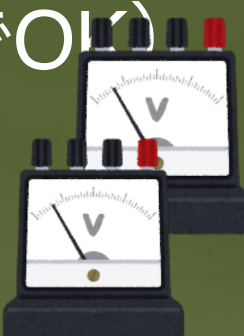
(施肥量以外の効果)

偶然誤差に転化

歪みはなくなり，結果を誤らない

無作為化の対象

- 順番を比較的容易に入れ替えられる要因は全てが無作為化の対象と考える
- 施肥量の例のように，無作為化するのは誤差要因（日照量）ではなく，目的要因でも良い（交絡している要因のいずれかでOK）
- 空間だけでなく，時間や作業者，測定器などの順番も重要

10.4　原則その３：局所管理

（系統誤差と偶然誤差を削減）

- 無作為化の限界①：製造ラインや収穫日，反復数が多い要因は順番を入れ替えるのが困難
- 無作為化の限界②：効果が大きな系統誤差を転化しても，F値の分母が大きくて検出が難しくなる

→こうした要因には局所管理を実施

→系統誤差をブロック因子として積極的に実験に取り込む方法（偶然誤差を減らす効果もある）

局所管理した圃場実験例

広い実験の場をブロック化

ブロック	ブロック2	ブロック3
施肥水準1	施肥水準1	施肥水準1
施肥水準2	施肥水準2	施肥水準2
施肥水準3	施肥水準3	施肥水準3

各ブロック内で実験を管理

各ブロックの中の日照量は同じ
（系統誤差がなくなる）

検定統計量

施肥量

ブロック因子
（日照量含む）

要因効果 ＋ 系統誤差

偶然誤差

F ←系統誤差が分離→F

こちらの検定が研究目的（精度が向上）

こちらの検定には興味なし

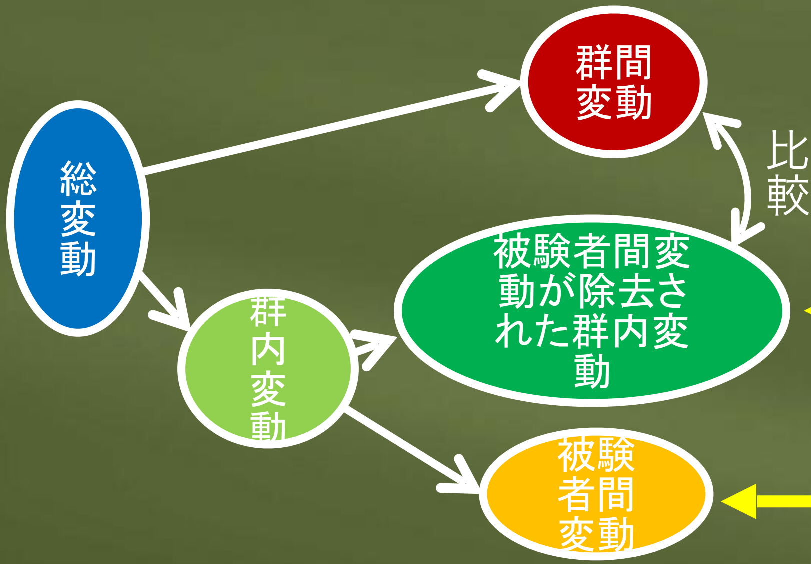
局所管理の長所（まとめ）

- 検定統計量から系統誤差が分離されるので，検定結果の正確さが向上する（系統誤差の効果が大きくても問題無し）
- 順番の入れ替え不要
- 局所の環境は均一なので，実験の場を全体で管理するよりも様々な偶然誤差が削減され，精度も向上する

局所管理した実験データの分析

対応のある一元配置（復習）

局所管理したデータは，ブロック因子も1つの要因と考え，対応のある一元配置か対応のない二元配置の分散分析を実施



ブロック差が除去された偶然誤差の変動

ブロック因子（系統誤差）による変動

ブロック化の対象例

- 官能検査：検査員
- 工場実験：製造ライン，原料ロット，日，作業者，出荷ロット，作業時間帯
- 農場実験：圃場の区画，植物工場の棚，果樹の個体，播種日，収穫日
- アンケート（聞き取り調査）：調査員，訪問地域，回答日

※注意：ブロック因子をたくさん導入すると実験や分析が複雑になるので，効果の大きい順に2つ程度に抑える

10.5　いろいろな実験配置

完全無作為化法と乱塊法

反　復
＋
無作為化

2つの原則を満たした実験
完全無作為化法
＋
局所管理

3原則全てを満たした実験
乱塊法（らんかいほう）

圃場全体で実験を管理

水準3	水準1	水準2
水準2	水準1	水準2
水準1	水準3	水準3

完全無作為化法

ブロック化

ブロック単位で実験を管理

水準3	水準1	水準2
水準2	水準2	水準1
水準1	水準3	水準3

乱塊法

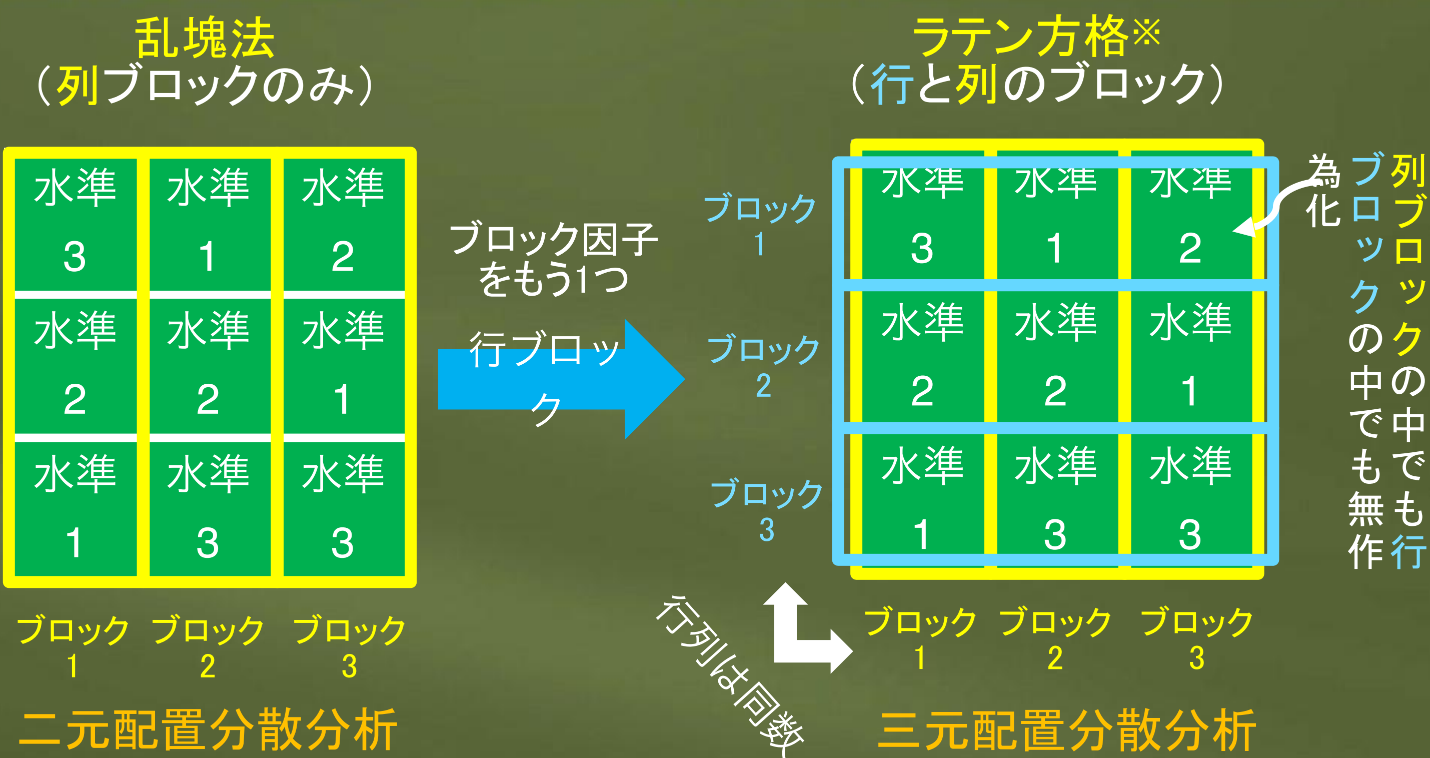
圃場全体で反復されていれば良い

乱塊法（局所管理を完全無作為化法に導入すること）の長所と短所

- 乱塊法は，無作為化だけでは対処できない規模の大きな実験や，個人差・実験日など大きな系統誤差が生じる可能性のあるときに向いている
- ブロック因子を導入することで，誤差分散の自由度をブロック因子と分け合うことで小さくなるため，系統誤差効果が小さい場合（や実験が小さい場合）には完全無作為化法でよい

ラテン方格法

ブロック因子をもう1つ導入したいときの実験配置



※注1：ブロック化の対象は空間に時間や人が混じっても良い。
2：昔は水準1・2にラテン語を当てていたことが名称の由来。

分割法

複数の実験要因を段階的に配置

- 検証したい要因が複数ある場合，全要因を同時に無作為化して配置するのは難しいことが多い
- 水準変更が難しい要因を一旦配置し，次に変更が容易な要因…と，段階的に分割して配置
- データは段階毎に分散分析にかけるべきだが，簡易的に多元配置分散分析を用いる

分割法の圃場実験例

灌水と施肥が収量に与える効果の検証

水準の入れ替えが面倒な方から

1次因子ａ：灌水（2水準）

a1	a2	a2	a1

変更1回目　変更2回目

完全無作為化法

（乱塊法やラテン方格法でもよい）

その後

必要に応じて3次，4次因子を設定

2次因子ｂ：施肥（3水準）

b1	b3	b2	b3
b3	b1	b1	b2
b2	b2	b3	b1

乱塊法

（1次の区画をブロックと見なす）

結局

2要因を同時に配置するよりも灌水水準の変更回数を節約（2回）できる

a1 b1	a2 b3	a2 b2	a1 b3
a1 b3	a2 b1	a2 b1	a1 b2
a1 b2	a2 b2	a2 b3	a1 b1

繰り返し2回の二元配置分散分析

10.6　直交計画法

実験の効率化

- 検証したい実験要因の候補がたくさんあると，とても大きな実験になってしまう

　→2水準の要因が7つの場合，2^7で128種類

- 限られた予算やスペースの下では，全実験の実施は困難

　→直交（配列）表を用いて，不要な実験を間引いて配置する（小さな実験で済ます）

直交（配列）表の例

ラテン方格　要因の数

$L_8(2^7)$

組合せ数　水準数

128種類から8種類に減数！

$L_8(2^7)$型直交表（たくさんある表でもっとも基本）

実験の種類（水準の組合せ）

列番号 / 実験番号	[1]	[2]	[3]	[4]	[5]	[6]	[7]
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

実験要因（誤差要因含む）

水準

水準ごとに4回反復しているので列の変動を計算できる

どの2列をとっても直交（相関係数＝0）

いろいろな直交表※

（巻末付録VIII〜Xを参照）

- 2水準系（全て2水準）：L_4, L_8, L_{16}など
- 3水準系（全て3水準）：L_9, L_{27}など

　→一部の交互作用が検定できる代わりに，要因を割り付ける列が決まっている（3列目は1・2列の交互作用など）

- 混合系（2水準と3水準が混在）：L_{18}, L_{36}など

　→交互作用の検定は全くできない（工学分野に適している）代わりに，要因を自由に割り付けられる

※注：直交表は自作することも可能です。

直交表の仕組み（L_4表を使って）

最小の2水準系L_4直交表

（普通は使わない）

$L_4(2^3)$	[1] 要因a	[2] 要因b	[3] 交互作用 a×b
実験①	1	1	1
実験②	1	2	2
実験③	2	1	2
実験④	2	2	1

なぜ3列目が1と2列目の交互作用なのか？

表頭に要因a，表側に要因bとした集計表

		要因a	
		水準1	水準2
要因b	水準1	実験①	実験③
	水準2	実験②	実験④

要因aとbの交互作用を見るには実験①・④と実験②・③のデータを比較する

要因をたくさん扱える理由と欠点

- 一部（あるいは全て）の交互作用をないものとして，水準の組合せを間引く

→交互作用がない場合や，あっても興味が無い場合（割付列に要注意）に有効

欠点①：間引いた組合せに大きな交互作用が隠れていても見つけられない

欠点②：反復数が少ないため精度は低い

L8表への要因の割り付けパターン

2水準の要因が3つの三元配置がスタート

列番号

[1]	[2]	[3]	[4]	[5]	[6]	[7]
a	b	a×b	c	a×c	b×c	a×b×c

要因3つと交互作用4つ
（三元配置実験）

興味の薄い交互作用に誤差を割り付ける

高次の交互作用はないので第4の要因dを割付

L8表の典型割付（次掲）

要因4つと交互作用2つ
（四元配置実験）

a	b	a×b	c	a×c	誤差	d

交互作用がないならば全列に要因を割付可

誤差がないと分散分析できない

要因6つ
（六元配置実験）

a	b	e	c	f	d	誤差

3列目が有意になっても要因eの効果ではなくa×bの交互作用かも知れないので注意（5列目も同様）

L8表の典型的な割付（四元配置実験）

主効果を確認したい要因の列だけを抜き出して実験を組む

（誤差や交互作用の列は計算に用いるだけなので無視）

	要因a	要因b	a×b	要因c	a×c	誤差	要因d
実験1	1	1	1	1	1	1	1
実験2	1	1	1	2	2	2	2
実験3	1	2	2	1	1	2	2
実験4	1	2	2	2	2	1	1
実験5	2	1	2	1	2	1	2
実験6	2	1	2	2	1	2	1
実験7	2	2	1	1	2	2	1
実験8	2	2	1	2	1	1	2

直交計画実験の分散分析

L8表の典型割付による四元配置実験とデータ

実験番号	[1] a	[2] b	[4] c	[7] d	データ
1	1	1	1	1	x_1
2	1	1	2	2	x_2
3	1	2	1	2	x_3
4	1	2	2	1	x_4
5	2	1	1	2	x_5
6	2	1	2	1	x_6
7	2	2	1	1	x_7
8	2	2	2	2	x_8
				総平均	

復習：群間変動は
反復数 × Σ（群平均−総平均）^{2}

（群平均$\bar{x}1$−総平均$\bar{\bar{x}}$）^{2}

これらの和に反復数（L8では4回）を乗ずれば群間変動になる

（群平均$\bar{x}2$−総平均$\bar{\bar{x}}$）^{2}

誤差変動も同じように計算し，それぞれ自由度（水準数−1）で割れば，その比が検定統計量F

分析ツールを使った直交実験の分析
（章末問題の問3より）

第10章末問題3から誤差列を除いたL8表

	a	b	a×b	c	a×c	d	データ
1	1	1	1	1	1	1	3
2	1	1	1	2	2	2	8
3	1	2	2	1	1	2	4
4	1	2	2	2	2	1	9
5	2	1	2	1	2	2	1
6	2	1	2	2	1	1	4
7	2	2	1	1	2	1	1
8	2	2	1	2	1	2	5

※注：3水準系や混合系にはダミー変数を使う

Excel分析ツールの回帰分析

Excel分析ツールの回帰分析の結果

	係数	標準誤差	t	P-値
切片	-0.12	0.93	-0.13	0.915
a	-3.25	0.25	-13.00	0.049
b	0.75	0.25	3.00	0.205
a×b	0.25	0.25	1.00	0.500
c	4.25	0.25	17.00	0.037
a×c	0.75	0.25	3.00	0.205
d	0.25	0.25	1.00	0.500

両側5％水準で有意

10.7　直交計画法の応用分野

- その1　品質工学におけるパラメータ設計：　安定した性能を発揮する製品を構成する要因・水準の組合せを（沢山の要因候補の中から）探るとき

- その2　マーケティングリサーチにおける仮想製品のプロファイル作成：　消費者が欲しいと思う製品の要因・水準をアンケートする際，回答者に提示する仮想製品の数を減らすとき

その1　品質工学（品質管理）

- 技術・新製品開発を効率的に行う技法体系

設計の最適化

パラメータ設計

工程の最適化

オンライン品質工学

システムの診断・予測

MTシステム

技術や製品を構成する要因・水準のたくさんの候補から直交表を使って品質の良い要因水準の組合せ（パラメータ）を探す

消費者が使用する様々な環境の下でも性能が安定して発揮されること
（たとえば，収量の高くなる栽培法を探すのではなく，異なる環境でも安定した収量が見込める栽培法を探す）

パラメータ設計の事例

L18直交表（交互作用を仮定しない混合系を使う）

検定は必須ではないので8列全てに要因を割付けても良い

制御の難しい外的要因の下での特性（計測）値

特性の安定性（測定値のバラツキの小ささ）を示す指標

※注：SN比にも数種類あり，いずれも複雑な計算が必要だが，ここでは分散の逆数（=1/VAR.P()）とする。

要因A	要因B	要因C	要因D	要因E	要因F	要因G	要因H	誤差1	誤差2	SN比
1	1	1	1	1	1	1	1	14.6	18.8	0.23
1	1	2	2	2	2	2	2	10.6	14.4	0.28
1	1	3	3	3	3	3	3	5.2	7.6	0.69
1	2	1	1	2	2	3	3	17.9	26.5	0.05
1	2	2	2	3	3	1	1	14.6	19.0	0.21
1	2	3	3	1	1	2	2	18.9	23.1	0.23
1	3	1	2	1	3	2	3	19.2	30.6	0.03
1	3	2	3	2	1	3	1	29.7	37.3	0.07
1	3	3	1	3	2	1	2	25.8	28.8	0.44
2	1	1	3	3	2	2	1	14.1	23.1	0.05
2	1	2	1	1	3	3	2	3.6	8.8	0.15
2	1	3	2	2	1	1	3	16.4	18.6	
2	2	1	2	3	1	3	2	26.1	34.7	
2	2	2	3	1	2	1	3	19.6	23.6	
2	2	3	1	2	3	2	1	12.4	14.0	
2	3	1	3	2	3	1	2	22.5	36.1	
2	3	2	1	3	1	2	3	31.5	39.9	
2	3	3	2	1	2	3	1	20.7	26.7	

↓分析ツールで回帰させた結果

（3水準は本来，ダミー変数を使

	係数	標準誤差	t	P-値
切片	-0.07	0.58	-0.11	0.91
要因A	0.09	0.16	0.59	0.57
要因B	-0.12	0.10	-1.26	0.24
要因C	0.29	0.10	2.91	0.02
要因D	-0.10	0.10	-1.00	0.34
要因E	0.04	0.10	0.43	0.67
要因F	0.10	0.10	1.02	0.33
要因G	-0.07	0.10	-0.72	0.49
要因H	-0.03	0.10	-0.27	0.80

要因Bを水準1，要因Cを水準3，要因Dを水準1，要因Fを水準3とする組み合わせが特性を安定させる

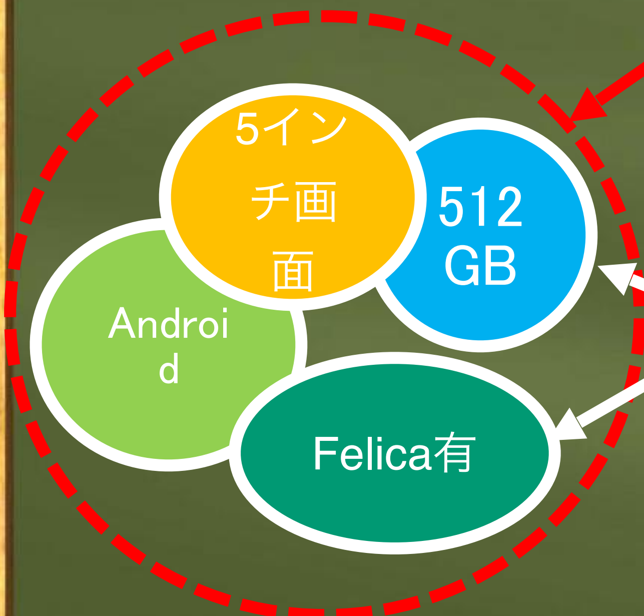
その2　コンジョイント分析
（マーケティングリサーチ）

- コンジョイント分析：消費者がどのような製品を好むのかを調べる市場調査（マーケティングリサーチ）のための手法
- 製品を構成する要因（属性）・水準の評価を個別に質問せず，店頭での買い物と同じように製品全体の評価を質問してデータを集める
- 調査時に示す製品プロファイル（属性・水準の組合せ）の種類が多いと回答者の負担になるために直交表を使って減数する

その2　コンジョイント分析

（マーケティングリサーチ）

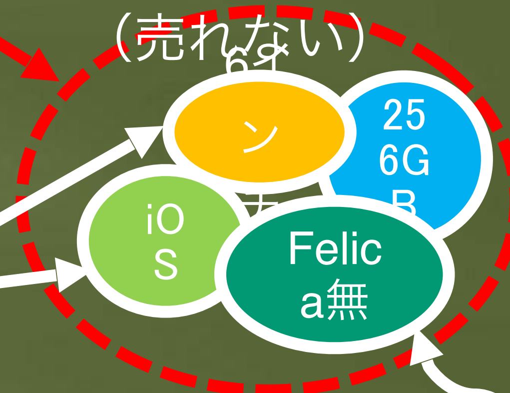
スマホの仮想製品A
（売れる）



消費者は総合的に評価

開発には部分評価が必要
（分析して把握）

スマホの仮想製品B
（売れない）



直交表で仮想製品の種類を減数

製品全体の効用は部分効用が結合したもの
※注：Conjoint＝結合

注：楕円の大きさは効用（使用満足度）の大きさ

コンジョイント分析の分析事例①

L8表を使って32種類から8種類に減数した仮想製品を評価させる

	画面 1:6インチ 2:5インチ	容量 1:256GB 2:512GB	Felica 1:無 2:有	OS 1:iOS 2: Android	評点
1	1	1	1	1	1
2	1	1	2	2	4
3	1	2	1	2	2
4	1	2	2	1	6
5	2	1	1	2	5
6	2	1	2	1	7
7	2	2	1	1	6
8	2	2	2	2	8

評点を回帰分析にかけて部分効用を計算

プロファイル
No.8
5インチ
212GB
Flica有
An

本当は32種類！

1番欲しいから8点（満点）！

プロファイルを示して評価してもらう

コンジョイント分析の分析事例②

属性を説明変数x，評点を被説明変数yとした重回帰分析の結果

属性	回帰係数
(定数項)	(-6.00)
画面	4.00
容量	0.50
Felica	2.00
OS	0.50

部分効用：（水準1を基準とした）水準2の評価

平均を0とした部分効用値に変換

部分効用のレンジが全体に占める割合

属性	水準	部分効用 -2 -1 0 +1 +2	重要度
画面	5インチ		57.1%
	6インチ		
容量	512GB		7.1%
	256GB		
Felica	有り		28.6%
	無し		
OS	Android		7.1%

もっとも重視される属性は画面サイズ

もっとも売れそうなスマホのスペック

以上で第10章は終了です。