

第 15 回 実際の統計解析

★ 教材「生物統計学__実験を始める前にすること 2013」を予習しながら空所を埋めておくこと

A. 実験を始める前にすること

1. 実験が始まる前にどのように統計解析を行うかを決めておこう

実験が終わり、データを取りおえてから、どのようにデータを解析するか、どんなグラフを書くかを決めればよいだろうというのではよくない。実験を始める場合、証明したい仮説があるはずである。その仮説をどのような図表で、どんな統計解析を使って、証明するかをあらかじめ決めておくことが実験計画をしっかり立てる上でも大切である。さらに実験データをとってから統計解析しようとする、適切な統計解析ができない実験計画だったことがあとからわかることも決して少なくない。

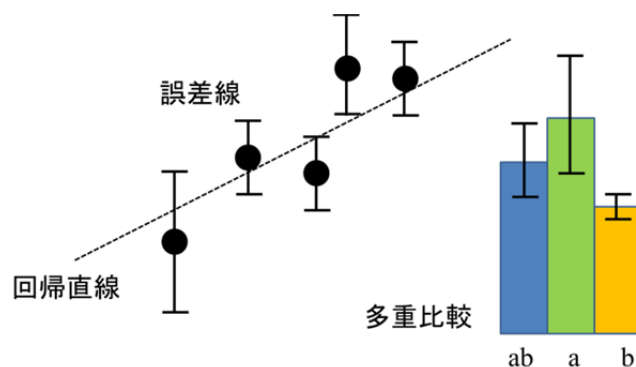
実験計画の手順

要因の絞り込み

- ① 要因を（ ）する
- ② 要因間の関係を（ ）する
- ③ 要因を（ ）する(制御因子、標示因子など)

因子と水準の決定

- ① （ ）する因子を決める
- ② それぞれの因子の（ ）を決める
- ③ 実験計画法に基づき実験計画を組む
- ④ どのような図表を書き、統計解析するかを決める



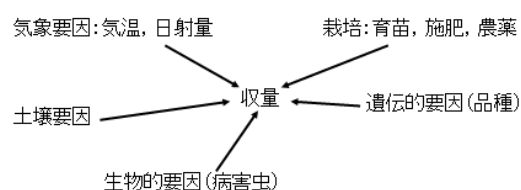
1. 要因（因子）をあげる

実験の目的が例えば、「人の血圧に及ぼす摂取塩分の影響」であっても、摂取塩分と血圧だけを測れば実験は終わりとは行かない。血圧に影響を及ぼす要因は多数あるから、摂取塩分の影響はそのような要因を無視しては解析できない。特に交互作用のある要因があればなおさらである。

したがって、まず実験目的が決まれば、目的とする結果に影響を及ぼしそうな要因をすべて挙げてみる。そのような要因を次に、①制御因子、②標示因子、③ブロック因子、④層別因子に分類する。

2. 要因間の関係を図示する

右の図はイネの収量に影響を及ぼす要因とそれぞれの関係を図示したものである。実際の実験では多数の要因のうち、少数のものに絞ってから、実験するが、実験を始める前に要因間の関係を図示しておくことは大切である。



3. 要因を分類する

① 制御因子

その最適条件（水準）を知るために取り上げる因子で，実験の間ではもとより，その結論を適用すべき（生産の）場においても，その条件を制御できるもの．

★ 制御因子を決める（複数でもよい）

- 1) 島根県で多収となる品種はどれかを定める 制御因子：品種
- 2) 多収となる品種とそれに適した作期を決める 制御因子：品種，作期
- 3) 多収となる品種とそれに適した作期・施肥量を決める 制御因子：品種，作期，施肥量

② 標示因子

その最適条件を知ることは直接の目的ではないが，この因子の水準が異なると，他の（制御）因子の最適条件が変わるおそれがある（**交互作用がある**）ために実験に取り上げる因子であって，**実験の間では制御されなければならないが，適用の間では必ずしも制御できない**．

- 1) 島根県で多収となる品種はどれかを定める 制御因子：品種

この場合，作期や施肥方法が標示因子となるかもしれない．

作期は水の得られる時期（梅雨など），水稻以外の作物（野菜，果樹，チャなど）の忙しい時期などによって，左右され，かならずしも現場の農家では制御できない．施肥方法でも，琵琶湖など湖沼，河川の近くのために水質保全の理由から，多収になる施肥方法が認められないこともある．

このように制御因子に交互作用のある標示因子が何かは専門的知識だけでなく，現場への理解も必要となることもある．

③ ブロック因子

実験の精度を高めるために，実験の間の局所管理に用いる因子で，その水準自身は特性値に若干の影響を与えるかもしれないが，他の（制御・標示）因子とは**交互作用を持たない**と考えられるもの．

例：水稻の品種試験においては圃場のムラなどである．

④ 層別因子

制御因子や標示因子と**交互作用を持つ**おそれがあるが，実験の間でも適用の間でも**制御できない**因子．

例：水稻の品種試験では，年度，地域などの因子である．同じ品種でも年によって成績が違ふこともあるし，地域によっても成績が異なるであろう．しかし，年度や地域は制御できない上に，品種との交互作用が認められる．

予習問題

自分の属する学科，講座，研究分野に関連するテーマで要因をできるだけあげ，そして因子の分類を考えてみよう．

① 要因をできるだけ考えてみよう

例) できるだけ環境負荷の少ない家庭排水を出すには？

天ぷら油の廃棄方法，米の研ぎ方，洗濯用の洗剤・・・

② ①で挙げた要因から 1 つ (複数でも可)，制御因子を決めよう．次に標示因子，ブロック因子，層別因子になりそうなものを考えてみよう．

制御因子

標示因子

ブロック因子

層別因子

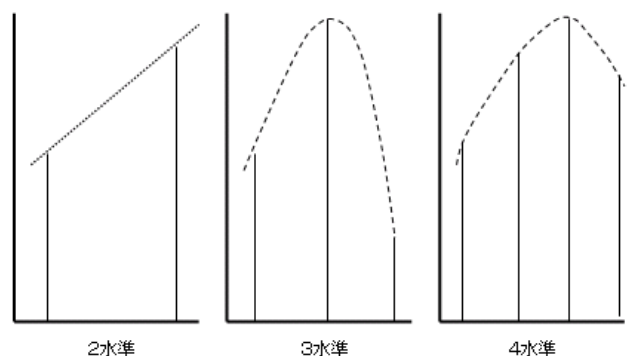
★ 教材「生物統計学__実験計画を立てる 2013」を予習しながら空所を埋めておくこと

4. 実験を計画する

次に実験を計画する．フィッシャーの 3 原則（反復・無作為化・局所管理）を満たすような乱塊法の実験計画が望ましい．実験計画が決まれば，どのような分散分析をするかが決まる．空欄の分散分析表をあらかじめ作っておくのもよい．

要因を複数同時に取り扱う要因実験の方がよい．複数の要因を同時に実験すると，① 実験の精度自体を高めること，② 交互作用を見積もれること，③ 実験のデータのばらつきをいろいろな角度から評価できる分散分析ができることなど有利な点が多い．

取り上げるべき実験の因子が決まれば，それぞれの因子についていくつ水準を取るかを定める．**質的因子**のときは利用できるあるいは考慮すべき数で決まる．質的因子とは例えば品種である．**量的因子**の場合，回帰分析することになる．直線を仮定するときは 3 つ程度，2 次曲線や対数，指数曲線を仮定するときは 4 つ程度の水準が必要である．水準はできるだけ広い範囲をカバーする方がよい．特に最初のうちから限られた範囲の水準でしか実験しないと，最適な値や新発見を逃すことになるかもしれない．実験による情報が増えれば，水準の幅を狭めていくことも可能である．



B. データを集めたら

1. 基礎統計量を計算する

① データの数の少ないとき

たいていの実験ではデータは 2~10 ぐらいであり、度数分布を調べるには少なすぎる。この場合は、平均、標準偏差、標準誤差を計算する。データが正規分布しないと平均、標準偏差はあまりデータの代表値としてふさわしくないこともある。さらに必要に応じて、平均値の区間推定などを行うこともできる。

② データの数の多いとき

データが 30 以上あれば、①で計算した平均、標準偏差などに加えて、度数分布、ヒストグラムを書いてデータの分布を調べることもできる。度数分布から、データが正規分布に近似できないと判断したら、メジアン、モードなども有用な統計量である。

2. () (外れ値) をチェックする

データの中に異常に大きい値、あるいは小さい値があるときの対処法は以下に述べるような方法がある。a. は必ず行うべきである。b. ~ e. はどれを用いるべきかは実験の目的、データの性質などを検討して、できる限り実験を開始する前に決めておく。

a. 異常値の原因が明らかなき

異常値の原因を調べ、測定におかしな点があるときは除去する。

b. 反復数を増やす

もし可能であれば、実験を繰り返し、データを増やすとそのデータが異常値であるかより明確になり、かつ異常値が平均値に及ぼす影響も軽減される。

c. 異常値の除去

異常値を除去したいときはスミルノフ法（スミルノフ・グラブス法）あるいはディクソンの Q テストで検定し、異常値であると認められるときは除去できる。根拠もなく、不都合なデータを捨てるのは好ましくない。

d. 内部平均（トリム平均）を用いる方法

集めたデータのうち両極端な値、すなわち最大値と最小値を除いたデータから得た平均を内部平均という。内部平均を用いるとデータのばらつきはかなり小さくなることが多い。データが 3 つのときの内部平均はメジアンと一致する。

e. ロバスト法を用いる方法

疑いのあるデータを完全には除去しないで、そのデータの重みを小さくすることによって、異常値の影響を軽減する方法がロバスト法である。適切なプログラムを利用しないとロバスト法は実行できない。しかし、異常値を除去することにもない問題（異常値であるかどうかを検定するときに p 値をどのように決めるのか、除去するかどうかで平均などが大きく変わってしまう）を回避できる。

3. 散布図を書き、データ間の相関を分析し、4. 分散分析を行う

得られたデータについてはどのデータとどのデータに相関があるかを散布図と相関係数を計算することで調べる。はじめは想定しなかった関係を見つけることもあるだろうし、最初に考えていたような関係がないこともあるだろう。データをグラフのように見える形にすることは相関関係を知るだけでなく、異常値やデータのグループ分けを見つける上でも重要なことである。実験計画で決めた分散分析をここで行う。

★ 教材「生物統計学__頻度データの推定と検定 2013」を予習しながら空所を埋めておくこと

C. 頻度（比率）の推定と検定

1. 頻度データの性質

着果率，発芽率，生存率のように全体のうち，何パーセントが着果，発芽，生存したかを示すデータを頻度データという。このような頻度データは等分散性を満たさず，さらに正規分布しないので，そのまま分散分析やt検定などをしない方がよい場合がある（頻度が30～70%の間に収まっている場合は，以下の数値変換はしなくてもよい）。

さらに頻度データの場合は調べた標本数が多くても，少なくともパーセント表示にしてしまうとその違いがなくなってしまうので，できるだけパーセントにしないで統計的検定をする方がよい。すなわちもし2つのデータがどちらも30%であったとしても，片方が100個体から得たデータ（すなわち30個体が該当した），片方が10個体から得たデータ（すなわち3個体が該当した）ではデータの意味が全然違うからである。

なお糖度，含水率のようなものはここでの頻度データには相当しない。糖度はパーセント表記でなくともg/100gのような表記も可能である（実際はこちらの表記が望ましい）。糖度などのように，もとの値が連続量である場合は頻度データに該当しないので，注意すること。

① 母比率の区間推定

母比率について区間推定する式はここでは説明しない。ここではデータをシートに代入すると計算できるようにした(エクセル授業用データ集2013の第15回頻度データの推定と検定タブ)。

2つの方法があり，簡便な方法が今でもよく使われる方法である。ふつうは簡便な方法を使って問題はない。より厳密な方法を使うこともコンピューターが普及した現在では増えている。

例題 25℃の飼育条件に置いたテントウムシの卵200個のうち，130個が孵化した。卵の母孵化率について95%信頼区間を付けて区間推定せよ。

右のように総数（卵の総数），対象数（ここでは孵化数），信頼率を入力すると，下に頻度（点推定値），区間推定値（上限と下限）が出力される。58.4%から71.6%が区間推定値であり，頻度データはサンプル数が200程度だと区間推定値の幅がかなりあることがわかる。より厳密な方法でも同じように計算できる。

よく使われる簡易な方法		
総数	200	
対象数	130	
信頼率	95 %	
頻度	65 %	
上限値	71.61034 %	
下限値	58.38966 %	
Z(α/2)	1.959964	

予習問題

30℃の条件でダイズの種子の発芽試験を行った。250 個の種子のうち、220 個が発芽した。95% 信頼区間を付けて母発芽率を区間推定せよ。（今回はよく使われる簡易な方法で計算すること）

② 母比率の検定

例題 A 商事が販売するカブトムシの幼虫について、A 商事は 75% がオス成虫になると主張している。この幼虫を 1000 匹購入したところ、720 匹がオス成虫になった。A 商事の主張の真偽について有意水準 5% で検定せよ。

帰無仮説 カブトムシの幼虫がオス成虫になる割合は 75% である

対立仮説 カブトムシの幼虫がオス成虫になる割合は 75% でない

右のように総数（幼虫の総数）、対象数（ここではオス成虫数）、仮説の比率（ここでは業者の主張する 75%）を入力すると、下に p 値が出力される。

母比率の検定	
総数	1000
対象数	720
仮説の比率	75 %
実際の比率	72 %
z 値	2.19089
p 値	0.02846

p 値は 0.028 なので、5% の有意水準で帰無仮説は棄却される。オス成虫になる割合は 75% でない。この場合、区間推定できるので、区間推定すると、オス成虫の出現割合は 95% の信頼率で 69.2~74.8% である。

よく使われる簡易な方法	
総数	1000
対象数	720
信頼率	95 %
頻度	72 %
上限値	74.78287 %
下限値	69.21713 %
Z(α/2)	1.959964

予習問題

現在、育種中の新しいミニトマトは従来の品種よりも裂果しないかもしれない。従来の品種の裂果率が 10% であった。新品種 200 個の果実を調査したところ、15 個が裂果した。新品種の裂果率は従来の品種と異なるかを 5% の有意水準で検定せよ。

帰無仮説

対立仮説

p 値

結論

③ 母比率の差の検定

例題 金魚に 2 種類の餌（餌 A と餌 B）を与えて、繁殖率を調べたところ、餌 A を与えた金魚 200 匹のうち、120 匹が繁殖に成功し、餌 B を与えた金魚 250 匹のうち、130 匹が繁殖に成功した。餌による繁殖率に差があるかを 5% の有意水準で検定せよ。

この検定はカイ二乗検定（分割表）でもできる。ここでは母比率の差の検定として検定してみよう。

帰無仮説 餌によって繁殖率は変わらない

対立仮説 餌によって繁殖率は変わる

右のように A の総数（餌 A を与えた金魚の総数）、A の対象数（餌 A を与えた金魚の繁殖数）、B の総数（餌 B を与えた金魚の総数）、B の対象数（餌 B を与えた金魚の繁殖数）を入力すると、下に p 値が出力される。

母比率の差の検定	
A の総数	200
A の対象数	120
B の総数	250
B の対象数	130
A の比率	60 %
B の比率	52 %
母比率	55.55556 %
z 値	1.697056
p 値	0.089686

p 値は 0.090 なので、5% の有意水準で帰無仮説は棄却されない。餌によって繁殖率が変わるとはいえないという結論になる。

予習問題

ある地方に絶滅が危惧されるカエルがいる。このカエルは温度によって雌雄の比率が決定するらしい。20 年前の調査では 200 匹のうち 105 匹が雄だった。今年の調査では 250 匹のうち 95 匹が雄だった。雌雄の比率が変化したかを有意水準 5% で検定せよ。

帰無仮説

対立仮説

p 値

結論

④ 頻度データ（パーセントデータ）を数値変換する方法

パーセント表記したデータを使って、分散分析、回帰分析をしたいときは数値変換してから、解析するのが望ましいことがある。

いくつかの手法のうち、ここでは、逆正弦変換を紹介する。ただしこの方法は、昔からよく使われた方法であるが、あまり近似の精度はよくない。しかし、計算は簡単なので、いまでも使われる。等分散性を修正するけれども、正規分布には近似しない。

全体の個数を n 、該当した個数を r とすると、頻度 $p=r/n$ である。頻度 p に対して、 $\arcsin(\sqrt{p})$ という変換を行う。 $r=0$ のとき（すなわち 0% のとき）は $p=1/4n$ 、 $r=n$ のとき（すなわち 100%）のときは $p=1-1/4n$ としてから $\arcsin$ 変換する。逆正弦変換の改良式も提唱されている (Sokal and Rohlf 著 Biometry など)。

近年では、ロジット変換の方がよく使われるようになっている。さらに頻度データを使った回帰分析（たとえば温度が上昇するとサクランボの着果率がどうなるかという問題が相当する）ではロジスティック回帰を利用することができる（ロジスティック回帰分析 丹後ら著、朝倉書店）。

★ 教材「生物統計学__一般線型モデル 2013」を予習しながら空所を埋めておくこと

D. 一般線形モデル（あるいは一般化線形モデル）

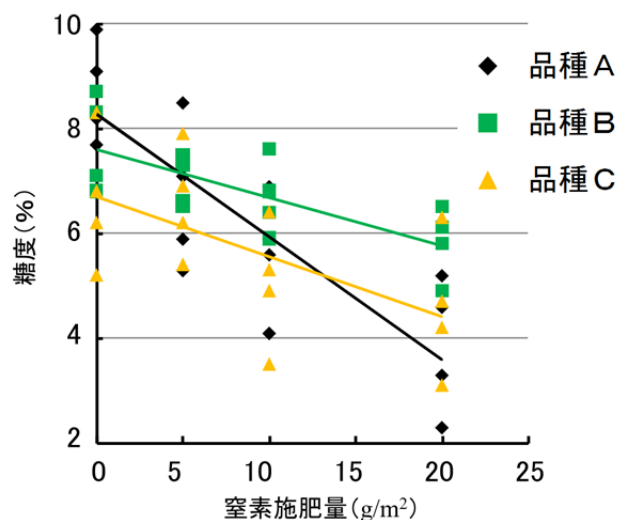
1. 一般線形モデル

これまでに学んだ種々の統計的手法は実際には同じ数学的枠組みの中で考えることができることがわかっていて、それらをまとめた統計的手法を一般線形モデル (general linear model, GLM) という。GLM を理解すると、 t 検定、分散分析、単回帰分析、重回帰分析、共分散分析などを 1 つのまとまった体系の中で利用できる。GLM の詳細は「一般線形モデルによる生物科学のための現代統計学-あなたの実験をどのように解析するか」、A. Grafen, R. Hails 著 共立出版を読むとよい。コンピューターの発達とそれに合わせた統計ソフトの充実で面倒な統計の計算を簡単にできるようになったために GLM の利用は現実的となった。GLM は因子が複数あり、しかも因子間に交互作用があるとき、とくに複数の因子において、量的因子と質的因子が混在しているときに有用な統計解析となります。

次のようなデータを考えよう。3 つの品種 (A, B, C) について、窒素施肥量とトマトの糖度の関係を調べたところ、以下のようなグラフの結果を得たとする。

グラフから、トマトの糖度は窒素施肥量に増加にともない、減少すること、その減少の程度（すなわち直線の傾き）は品種によって異なること、品種 B は品種 C より糖度が高いことがいえそうである。

しかし、従来型の繰り返しのある二元配置分散分析を行うと、交互作用の p 値は 0.114 となり、5% の有意水準では交互作用は認められなくなる。これ



は要因としてとりあげた窒素施肥量を量的因子（共変量ともいう）ではなく、品種と同じ質的因子（カテゴリカル因子ともいう）扱いして、分散分析したためである。GLMでは因子をそれぞれに質的因子（カテゴリカル因子）、量的因子（共変量）を指定できる。その結果、GLMによる分散分析では交互作用のp値は0.0096となり、1%の有意水準で交互作用が認められる。

GLMではさらに品種Aについて窒素施肥量が7g/m²のときの糖度はいくらかという推定・予測も可能である。従来の分散分析では統計的検定をするだけだったのに対し、統計的推定・予測を統計モデルを作って可能にする点でもGLMは有用である。さらにGLMを一般化した一般化線形モデルはデータがそのままでは正規分布しない場合でも統計解析をしやすくなる。

2. 分散分析と誤差の制御

① 質的因子と量的因子の区別

GLMのところで紹介した実験データを解析してみよう。

窒素施肥量とトマトの糖度の間には関係がありそうだとA君は考えた。とりあえず園芸店に行き、トマトを16個購入し、窒素施肥量4段階（0, 5, 10, 20g/m²）の処理を与え、トマトの糖度を測定した。すると右の表のような結果を得た。

A君は何も考えずに窒素施肥量という因子が1つある分散分析すればよいと次のように分散分析し、多重検定までしてしまった。

施肥量(g/m ²)	糖度(%)
0	6.1
0	7.3
0	8.3
0	9.2
5	5.7
5	6.5
5	7.2
5	7.9
10	4.2
10	5.6
10	6.3
10	6.7
20	2.9
20	4.5
20	5.3
20	6.3

分散分析表

変動要因	変動	自由度	分散	割られた分散	P-値	F境界値
グループ間	20.235	3	6.745	4.545914	0.023836	3.490295
グループ内	17.805	12	1.48375			
合計	38.04	15				

分散分析の結果、5%の有意水準で窒素施肥量によってトマトの糖度は変化することがわかった。でも施肥量0g/m²と20g/m²の間にしか有意差がないのか・・・とA君は判断した。20g/m²も施肥を変えないと糖度に有意差が出ないのか、施肥の効果はたいしたことないな！とA君は思った。

施肥量(g/m ²)	糖度(%)
0	7.725 a*
5	6.825 a, b
10	5.7 a, b
20	4.75 b

* 同一の符号間では有意水準5%でTukey HSD testによって有意差がない

これをみたS君はあぜんとして、なぜ施肥量を量的因子（共変量）にしないのかといった。施肥量を共変量として分散分析、すなわち回帰分析すればもっと精度よく解析できるはずだ。

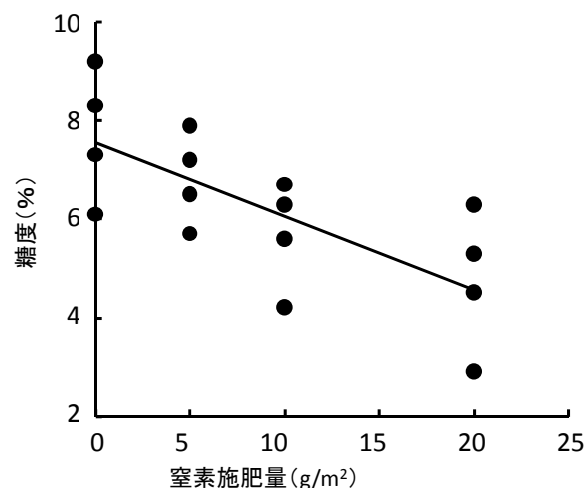
S君の指導にしたがって、回帰分析したら以下の結果を得た。

分散分析表

	自由度	変動	分散	割られた分散	有意 F
回帰	1	19.46314	19.46314	14.66793	0.001839
残差	14	18.57686	1.326918		
合計	15	38.04			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 99.0%	上限 99.0%
切片	7.555	0.446136	16.93428	1.01E-10	6.598132	8.511868	6.226922	8.883078
施肥量(g/	-0.14914	0.038942	-3.82987	0.001839	-0.23267	-0.06562	-0.26507	-0.03322

p 値はさっきの 0.024 に比べて、0.002 と 10 分の 1 になり、1%の有意水準でトマトの糖度に対する窒素施肥量の効果が認められた。しかも回帰直線の傾きは -0.149 だから施肥量が 1g/m^2 増えると糖度は平均して、 0.149% 低下することもわかった（右のグラフ）。



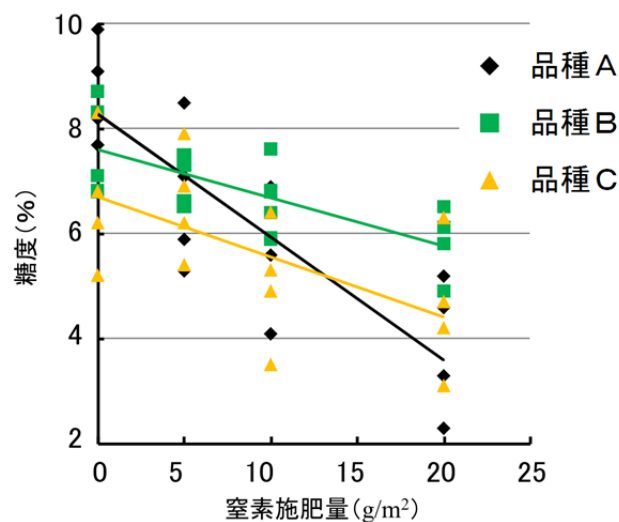
② 品種の違いを無視していたなんて・・・交互作用の発見

S 君はA君のデータをよくみると複数の品種をごちゃ混ぜにしていることに気づいた。A君はどのデータがどの品種になるかもきちんと記録していたのだ。まめに詳細なデータをとっていることは感心だ。でもA君は自分は窒素施肥量とトマトの糖度の関係を知りたいだけだから、解析のときに品種まで考える必要はないのでは主張した。

S 君はたとえ窒素施肥量の効果にしか関心がなくても、現場の農家は特定の品種しか栽培しないこともあるだろう。例えば、その品種がその地域のブランドであるとか、その品種の栽培になれているとか…そういう場合、品種と窒素施肥量の間に交互作用があると、品種を標示因子として取り込む方がいいかもしれない。

S 君は交互作用があるかないかはグラフを書いてみるのがわかりやすいと指摘した。A君は各品種で 4 反復、品種数は 3 つの実験をもう一度行い、右のようなグラフを書いた。

交互作用があるようだ（どの品種に交互作用があるだろうか・・・）



予習問題

右上のグラフをみて、窒素施肥量と品種の間に交互作用があることがわかった。すなわち窒素施肥量と糖度の関係は品種によって異なるということである。ではどの品種に交互作用が強くみられたかを考えよ。

GLM で分散分析した結果，以下のような結果を得た．

変動因	自由度	平方和	分散	F 値	p 値
品種	2	9.255	4.628	4.395	0.018
窒素施肥量	1	57.239	57.239	54.365	0.000
品種×窒素施肥量	2	10.947	5.474	5.199	0.0096
誤差	42	44.220	1.053		

品種の効果は 5%の有意水準で認められ，窒素施肥量の効果は 0.1%の有意水準で認められた．しかも品種と窒素施肥量の間には交互作用が有意水準 1%で認められた．

S 君はもし品種 A しか栽培しない農家と品種 C しか栽培しない農家がいたら同じ窒素施肥量でも全然違う糖度になるかもしれないんだと説明した．

③ ブロック因子によってさらに精度をあげることができる

S 君は A 君が実験している大学農場を訪れた．ガラス室が 20 棟もある立派な大学農場だ．しかし，ガラス室の日当たり，中の温度，土壌が微妙に違うようだ．どうもガラス室によって多少の生育の差があるらしい．でも品種や糖度の間には交互作用はなさそうだ．

S 君は A 君にガラス室 1 つを 1 つのブロックとして，1 つのガラス室内はできるだけ均一に管理する局所管理でガラス室による系統誤差を除去するのがよいとアドバイスした．実は②のデータは異なる 4 つのガラス室に各品種をそれぞれ栽培して得たデータだった．

GLM で分散分析した結果，以下のような結果を得た．

変動因	自由度	平方和	分散	F 値	p 値
品種	2	9.255	4.628	5.814	0.006
窒素施肥量	1	57.239	57.239	71.907	0.000
ガラス室	3	13.176	4.392	5.517	0.003
品種×窒素施肥量	2	10.947	5.474	6.876	0.003
誤差	39	31.045	0.796		

品種の効果は 1%の有意水準で認められ，窒素施肥量の効果は 0.1%の有意水準で認められた．しかも品種と窒素施肥量の間には交互作用が有意水準 1%で認められた．ガラス室の効果すなわちブロックの効果も 1%の有意水準で認められた．

S 君はこれだけ小さな有意水準でも有意差があるなら，この実験の結果は信頼してよいとみんな考えてくれるだろうといった．

E. 有意水準の意味（5%，1%水準で有意であるとはどういうことか？）

統計的に5%水準で有意であるとしても、実際問題としてその差に意味があるかは別の問題である。例えば、3種類の飼料A、B、Cでニワトリの産卵量を調べた結果、ニワトリを何千羽も使った結果、わずかに0.01個の産卵量の差であるが、Aが有意に産卵量を増加させたという結果が得られたとしよう。0.01個の産卵量の差は実際問題としては差がないといえるが、統計的には供試数を増やせば有意差は出やすくなる。有意であるというのは、差がないという帰無仮説を指定した確率の下で棄却できるということを示すだけである。

つまり、1%水準で有意差のあるデータは5%水準で有意差のあるデータよりも意味があるということではなく、ただ統計的に母平均に差のあることをより高い確信を持っていえるというだけである。統計的に有意であるという結果を得たら、つぎにその差が現実の意味があるかを判断しなければならないが、このためには専門的な知識を動員する必要がある。なお、逆に平均の差が非常に大きく、現実的に意味がありそうでも、統計的に有意でない場合は、差があるとはいえないとすべきである。この場合は、証拠が少ないということであり、さらにデータを増やすべきである。

5%水準では有意であるが、1%水準では有意ではないという場合はどう考えるか？この場合は実験の目的によって判断が変わるであろう。医薬品のように間違いのある確率が低くあるべきものは5%水準で有意であるぐらいでは採用せずに、さらに証拠を増やし、1%水準で有意であるようにすべきだろう。一方、工場で製品を作るときはあまりに慎重だと利益を失うと判断するのなら、5%水準で有意であれば採用すべきだろう。もしかすると10%水準でも採用すべきかもしれない。このように有意水準は帰無仮説を棄却できる確率を示すにすぎない。要するにある新しい技術を採用したらコストがいくら下がるという利益は実験を繰り返してより低い有意水準としたところで大きくなるわけではない。あるいはサツマイモの品種Aと品種Bの収量を比較した結果、5反復の実験で5%の有意水準で収量に有意な差が認められた場合と20反復の実験をやった結果、0.1%の有意水準で収量に有意差が認められた場合で、後者の方が品種Aと品種Bの収量の差が大きくなるということはない。すなわち実際上の収量の差は有意水準とは関係がない。

G. 定期試験について

日時 2月12日（水） 午後12時45分から2時15分まで（90分間）

場所 マルチメディア演習室1

すべて持ち込み可。パソコンは必ず持ってくる。ただし通信機能の使用は不可。

問題の入ったエクセルファイルをあらかじめMoodleからダウンロードしてください（2月7日ごろMoodleに載せる予定）。問題の入ったファイルはパスワードを入力しないと開けられません。パスワードは試験当日に発表します。答えは紙に書きます。Moodleでファイルを提出する必要はありません。

出題範囲 全部

成績評価：定期試験＋レポート（宿題と予習）＋授業での演習問題など

宿題未提出・再提出分の最終締め切り 2月11日（火）午後5時

宿題を期限までに全部出せば、期末試験の結果が悪くても追試を必ず受けることができます。