

第14回 実験計画と統計解析の実際

A. 多変量解析

1. 多変量解析とは？

現実の現象は1つだけの要因だけに支配され、それから説明されるということはむしろまれなことである。相関分析や単回帰分析では2つの変量間の関係を調べる。しかし、3つ以上の変量間の関係を調べる必要のあることは多い。このような多数の変量間の関係を調べる方法をまとめて多変量解析という。多変量解析には目的に応じて、多数の手法がある。ここではその一つである重回帰分析を紹介する。

2. 重回帰分析

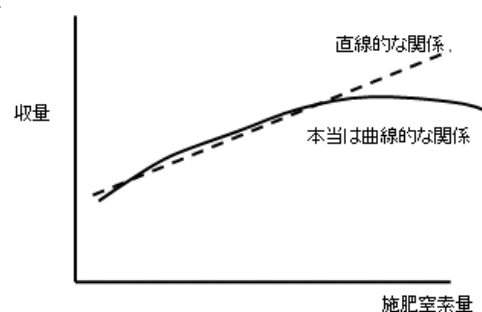
先週の講義でとりあげた果実の糖度の例で考えよう。果実の糖度は実際には平均気温以外に日射量、土壌水分、肥料などさまざまな要因で決まると考えられる。目的変数である果実の糖度に対して、説明変数（独立変数）を2つ以上考えたいということがある。単回帰分析は1つの説明変数であったが、これを複数に拡張したものが重回帰分析である。

説明変数間に相関があるとき、単回帰分析の式を足しあわせただけでは正しい重回帰式は得られない。例えば、平均気温と日射量にはおそらくかなり強い正の相関があるだろう。重回帰分析を行うとそのような説明変数間の相関を除いて、各説明変数単独の効果を評価できる。

3. 重回帰分析を使う上での注意

重回帰分析は単回帰分析の拡張であるから、以下述べる回帰分析の注意はおおむね当てはまる。

説明変数の範囲より外ではあまり精度がよくない。回帰分析では得られたデータの範囲外のことはあまり精度が高くない。さらに実は曲線で当てはめるのがよいデータでも部分的には直線がよく当てはまる場合もあり、この場合、データの範囲外ではきわめて精度が悪くなる。トレンド解析のように時間の経過にともなうデータの解析は、そのデータの本質上、データの範囲外を予測する回帰分析となるから、このように精度が悪くなるということを留意して用いなければならない。



重回帰分析でも同じことがいえる。得られたデータの範囲より外の説明変数について、目的変数を計算することはあまりよくない。例えば、20~30°Cの平均気温と15-25MJ m⁻² d⁻¹の日射量から糖度を説明する重回帰式を作ったとしよう。この場合、平均気温15°Cの場合かどうか、あるいは日射量5MJ m⁻² d⁻¹のどうかということはあまり精度よく推定・予測できない。

4. 単回帰分析と重回帰分析の適用現場の違い

重回帰分析はある目的変数を説明する要因がいくつも考えられるときに、説明変数を絞り込むために用い、単回帰分析は絞り込んだ説明変数が他の重要な要因は一定になるように制御された条件で目的変数にどのように影響するかを定量化する場合に用いることが多いようである。

B. 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 (しかし、なぜか農学分野ではいまだに利用が少ない)。

GLM の背景にある数学を理解する必要はなく、分散分析と単回帰分析の2つの原理をしっかりと理解すれば、GLM の原理を理解することは難しくない。データの変動を要因 (さらに要因間の交互作用) ごとに分けていくだけである。GLM は要因が複数あり、しかも要因間に交互作用があるとき、とくに要因が変数であるときに強力な威力を発揮する。

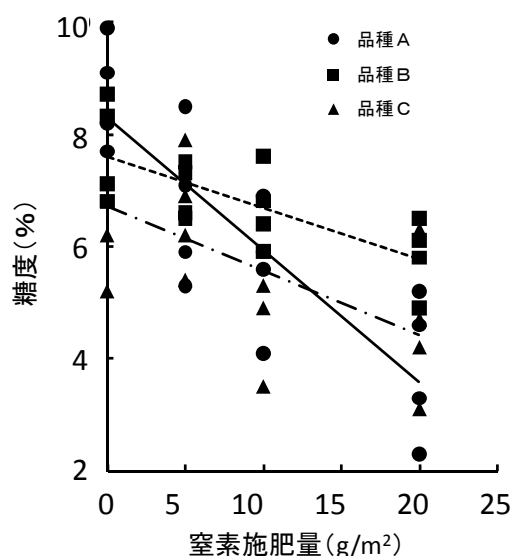
次のようなデータを考えよう。3つの品種 (A, B, C) について、窒素施肥量とトマトの糖度の関係を調べたところ、以下のようなグラフの結果を得たとする。

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変数因子ではなく、品種と同じ質的因子扱いして、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、変数因子 (正確には共変数因子という) 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/m^2 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容易にする

GLM の利用は島根大学では総合情報処理センターにおいて SPSS を利用することによって可能である。また、多数の統計ソフトが GLM に対応しているので、各自の研究室で利用している統計ソフトを利用することも可能である。

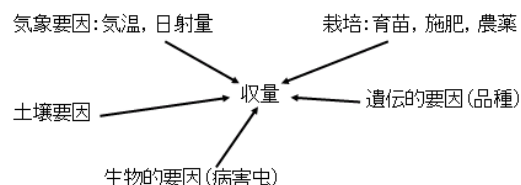


C. 実験を始める前にすること

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決める 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決める 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）、水稲以外の作物（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、実験の場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、実験の場でも適用の場でも**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うこともあるし、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。

練習 1

次の文章を読んで、どれが制御因子、標示因子、ブロック因子、層別因子かを考えよ。

湾内でカキの養殖をするための実験を考えよう。養殖業者として選択できるのは餌の種類と稚貝の採取場所の2つだとする。水産試験場ではさらにカキの種類を選べる。カキの種類と餌の種類には交互作用があるらしい。いかだを置く水深によって、カキの成長が異なるが、水深は他の要因との間に交互作用はない。なぜか年度によって、カキの成長は異なり、しかも稚貝の採取場所と交互作用がみられる。

制御因子

標示因子

ブロック因子

層別因子

練習 2 乳牛の泌乳量を最適にしたい。

1) 考えつく限りの要因をあげてみよう。

2) 1) で挙げた要因のうち、どれか一つを制御因子として、上であげた因子がそれぞれ標示因子、ブロック因子、層別因子になるかを考えてみよう。

制御因子

標示因子

ブロック因子

層別因子

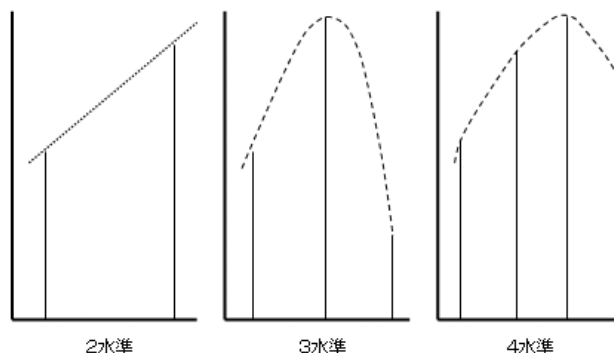
他の例を自分の属する講座，研究分野に関連するテーマで考えてみよう。

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。



5. どのような統計解析をするかを定める

データを取り終わってからどういう解析をするかを定めるのは本当は正しくない.

2つの母集団の平均値の有意差検定 t 検定

3つ以上の母集団の平均値の有意差検定 分散分析

分散分析の後, 質的因子の場合は多重検定, 量的因子の場合は回帰分析を行い, 最適な水準がどれかを定めることができる.

練習: どの統計解析法を用いるかを () に入れよ. () に入れる語句は, t 検定, 一元配置の分散分析, 二元配置の分散分析, 回帰分析, 多重検定とする.

おいしいパンを作る実験をする. パンの味はパンの味センサーというもので量的に定量的に測ることができるとする.

① 発酵の有無で味がどうなるかを比較実験する. () で検定する.

② 発酵の温度 (3段階: 15,25,35°C) による味の違いを比較実験した. () で3つの水準の平均には差がないという帰無仮説を検定する. 帰無仮説が棄却されたなら, 発酵の温度によって味が変わることが示されたので, 次は最適の発酵温度を () で求めることができる.

③ イースト菌の種類 (3種類: 酵母 a,b,c) による味の違いを比較実験した. () で3つの水準の平均には差がないという帰無仮説を検定する. 帰無仮説が棄却されたなら, 発酵に使う酵母の種類によって味が変わることが示されたので, 次は最適の酵母を () で決めることができる.

④ 発酵の温度 (3段階: 15,25,35°C) とイースト菌の種類 (3種類: 酵母 a,b,c) の組み合わせによる味の違いを比較実験した. () で発酵の温度とイースト菌の種類がそれぞれ味に影響するかを検定できるだけでなく, 両因子間の交互作用も検定できる.

D. データを集めたら

1. 基礎統計量を計算する

① データの数が多きとき

データが30以上あれば, 度数分布, ヒストグラムを書いてデータの分布を調べる. 次に平均, 分散, 標準偏差などを計算する. 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 データの数の少ないとき

たいていの実験ではデータは5ないし10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、分散、標準偏差を計算する。データが正規分布しないと平均、分散はあまりデータの代表値としてふさわしくないので、実験を始める前に確認しておくこと。

③ 異常値をチェックする

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. ～d. 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。

2. 母平均を95%信頼区間などをつけて推定する

自分の得たデータを平均だけみて判断するのは好ましくない。どの程度データの平均が信頼できるかの指標である標準誤差を計算しておくのがよい。データのばらつきが自分の求める精度よりも大きいときは、ばらつきを小さくすることを検討する。

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。

E. 分散分析の結果をどう読むか？

1. 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ならば、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。

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って、トマトを16個体購入し、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	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 (statistics : 統計の s だ) はあぜんとして、なぜ施肥量を量的因子 (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。

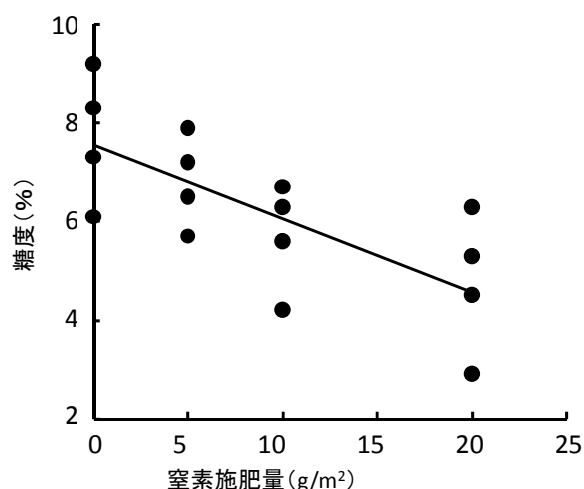
S君の指導にしたがって、回帰分析したら以下の結果を得た。

分散分析表

	自由度	変動	分散	割られた分	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/m² 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)。



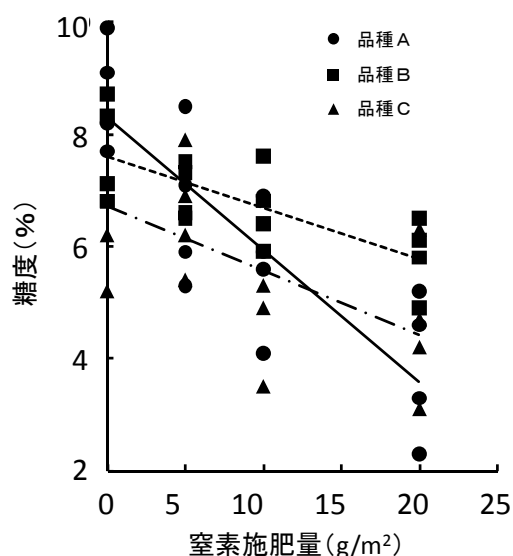
② 品種の違いを無視していたなんて・・・交互作用の発見

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・・・）



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。

F. 定期試験について

日時 2月3日(火) 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実験計画学のホームページからダウンロードしてください(1月31日ごろホームページに載せる予定)。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。フロッピーやCDRを用意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+授業中の質問

レポート未提出分の最終締め切り 1月30日(金)午後5時

締め切りまでに再提出が残った場合は、追試を受ける資格を失います。