実験計画学
2005 年 1 月 25 日

# 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？

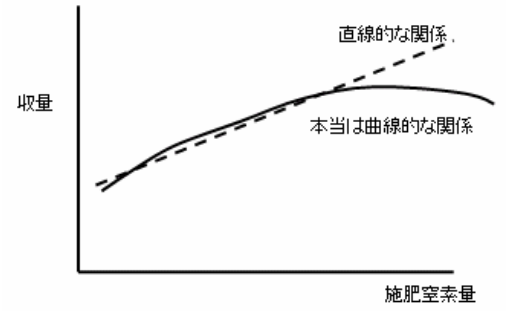
現実の現象は１つだけの要因だけに支配され，それから説明されるということはむしろまれなことである．相関分析や単回帰分析では２つの変量間の関係を調べる．しかし，３つ以上の変量間の関係を調べる必要のあることは多い．このような多数の変量間の関係を調べる方法をまとめて多変量解析という．多変量解析には目的に応じて，多数の手法がある．ここではその一つである重回帰分析を紹介する．

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より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，データ外ではきわめて精度が悪くなる．

トレンドのように時間がたつとどうなるかというデータは，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，このような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℃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るとしよう．この場合，平均気温 15℃の場合はどうか，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．

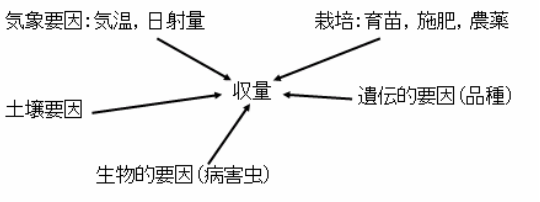
**B．実験を始める前にすること**

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因子のうち，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，実験の場ではもとより，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，**実験の場では制御されなければならない**が，適用の場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：水稲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：水稲の品種試験では，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：乳牛の泌乳量は何で決まるか.

１）考えつく限りの要因をあげてみよう.

２）上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.

標示因子

ブロック因子

層別因子

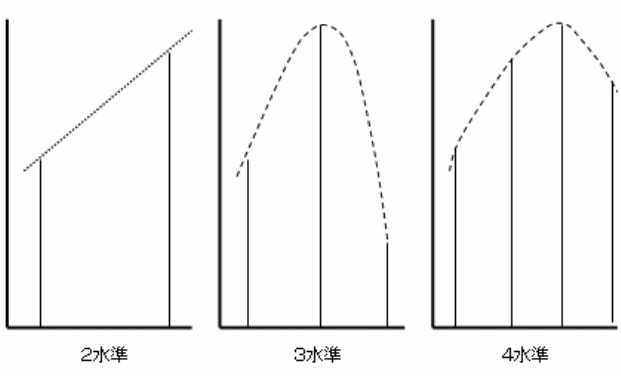
４．実験を計画する

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決める．**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決める

データを取り終わってからどういう解析をするかを決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決めることができる．

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℃）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℃）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．

## C．データを集めたら

１．基礎統計量を計算する

① データの数が多いとき

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る．

ａ．異常値の原因が明らかなとき

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．

③　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．

２．データ間の相関を分析する

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．

３．目的に応じた統計解析を行う

実験計画で決めた分散分析をここで行う．

**Ｄ．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に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，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．

２．分散分析表を読む

①　一元配置分散分析表

次のようなデータを一元配置の分散分析した結果を解釈してみよう.

水稲の1穂穎花数の品種間差異を5品種について調査した。

| | コシヒカリ | 日本晴 | NPT160 | IR72 | タカナリ |
|---|---|---|---|---|---|
| | 56 | 80 | 188 | 131 | 177 |
| | 52 | 84 | 169 | 129 | 207 |
| | 64 | 64 | 131 | 106 | 179 |
| | 60 | 71 | 233 | 95 | 173 |
| | 17 | 81 | 186 | 152 | 109 |
| | 98 | 93 | 183 | 152 | 151 |
| | 122 | 107 | 167 | 145 | 143 |
| | 101 | 104 | 135 | 144 | 162 |
| | 128 | 77 | 148 | 153 | 123 |
| | 85 | 60 | 126 | 83 | 188 |
| | 125 | 96 | 177 | 160 | 96 |
| | 60 | 111 | 158 | 115 | 152 |
| | 108 | 78 | 127 | 113 | 162 |
| | 85 | 80 | 162 | 71 | 111 |
| | 83 | 80 | 117 | 76 | 139 |
| | 97 | 94 | 144 | 101 | 154 |
| | 84 | 85 | 165 | 76 | 132 |
| | 108 | 92 | 98 | 68 | 214 |
| 平均 | 85.16667 | 85.38889 | 156.3333 | 115 | 154 |
| 標準偏差 | 29.40038 | 13.9332 | 31.93007 | 31.8932 | 32.76476 |

分散分析: 一元配置

概要

| ｸﾞﾙｰﾌﾟ | 標本数 | 合計 | 平均 | 分散 |
|---|---|---|---|---|
| コシヒカリ | 18 | 1533 | 85.16667 | 864.3823529 |
| 日本晴 | 18 | 1537 | 85.38889 | 194.1339869 |
| NPT160 | 18 | 2814 | 156.3333 | 1019.529412 |
| IR72 | 18 | 2070 | 115 | 1017.176471 |
| タカナリ | 18 | 2772 | 154 | 1073.529412 |

分散分析表

| 変動要因 | 変動 | 自由度 | 分散 | 観測された分散比 | P-値 | F 境界値 |
|---|---|---|---|---|---|---|
| ｸﾞﾙｰﾌﾟ間 | 88362.38 | 4 | 22090.59 | 26.49545522 | 2.81E-14 | 2.479013 |
| ｸﾞﾙｰﾌﾟ内 | 70868.78 | 85 | 833.7503 | | | |
| 合計 | 159231.2 | 89 | | | | |

変動要因のグループ間とは処理による変動（ここでは品種のちがいによる変動）を指し，グループ内は誤差変動を指す. 処理による変動の分散は誤差分散に比べて十分に大きいことがわかる. エクセルの出力した表においてＰ値が 0.05 より小さければ５％水準で, 0.01 より小さければ１％水準で有意である.

上のエクセルの出力を分散分析表に直すと下のようになる．論文ではたいてい下のような形式で分散分析の結果を示してある.

| 要因 | 自由度 | 平方和 | 平均平方（分散） | 分散比（Ｆ値） |
|---|---|---|---|---|
| 全体 | 89 | 159231.2 | | |
| 品種 | 4 | 88362.38 | 22090.59 | 26.495*** |
| 誤差 | 85 | 70868.78 | 833.750 | |

分散比（Ｆ値）の右肩に有意差があることを示す記号をつける慣習がある. *が 1 つのときは 5%水準で有意であることを，2 つのときは 1%水準で有意であることを，3 つのときは 0.1%水準で有意であることを示す場合が多い．最近はＰ値そのものを掲載する場合も増えている.

② 繰り返しのない二元配置の分散分析表

第９回分散分析その２p9 の例を以下に再掲した.

ヤギに与えると成長がよくなるという5種類の薬品(対照区を含む)とふだんの餌5種類との二元配置の分散分析

| | 麦わら | 稲わら | 牧草 | 濃厚飼料 | 雑草 |
|---|---|---|---|---|---|
| 対照区 | 11 | 5 | -1 | -4 | 2 |
| A | 29 | 17 | 14 | 2 | 20 |
| B | 8 | 14 | 20 | 20 | 26 |
| C | 23 | 8 | 8 | 11 | 5 |
| D | 26 | 20 | 11 | 5 | 17 |

分散分析表

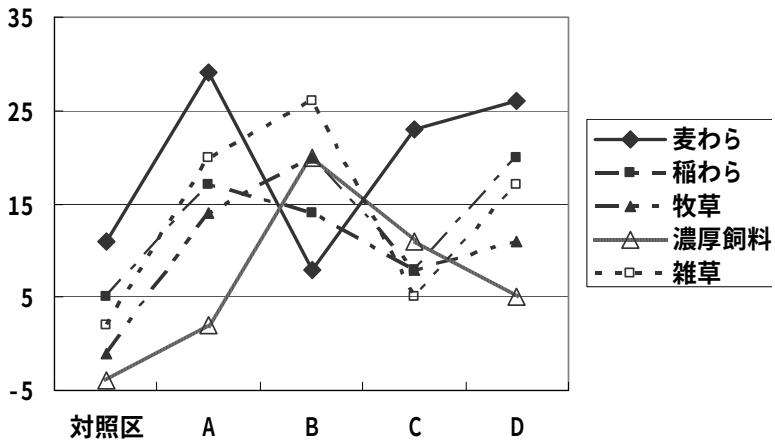
| 変動要因 | 変動 | 自由度 | 分散 | 観測された分散比 | P-値 | F 境界値 |
|---|---|---|---|---|---|---|
| 行 | 761.04 | 4 | 190.26 | 4.15324165 | 0.017033 | 3.006917 |
| 列 | 433.44 | 4 | 108.36 | 2.365422397 | 0.096597 | 3.006917 |
| 誤差 | 732.96 | 16 | 45.81 | | | |
| | | | | | | |
| 合計 | 1927.44 | 24 | | | | |

行はここでは薬を表し，P値が 0.05 より小さいので，薬は 5%の有意水準で有意であり，薬の効果が認められた．列はここでは餌の違いを表し，P値が 0.05 より大きいので，餌の違いによる効果は 5%の有意水準で認められなかった．繰り返しのない二元配置では交互作用は検出できない.

以上のようなエクセルの出力をふつうの分散分析表に直すと，以下のようになる．ns は 5%水準で有意差がないことを示すことが多い．ns は英語の not significant の略である.

| 要因（処理） | 平方和Ｓ | 自由度φ | 平均平方（分散） | 分散比（Ｆ値） |
|---|---|---|---|---|
| 薬 | 761.04 | 4 | 190.26 | 4.15* |
| 餌 | 433.44 | 4 | 108.36 | $2.37^{ns}$ |
| 誤差 | 732.96 | 16 | 45.81 | |
| 合計 | 1927.44 | 24 | | |

しかし，グラフを書いてみると交互作用があるようだ（下の図）.



③　繰り返しのある二元配置の分散分析表

第 9 回分散分析その 2 p10 の例を以下に再掲した.

ヤギに与えると成長がよくなるという5種類の薬品(対照区を含む)とふだんの餌5種類との二元配置の分散分析

| | 麦わら | 稲わら | 牧草 | 濃厚飼料 | 雑草 |
|---|---|---|---|---|---|
| 対照区 | 5 | 4 | 0 | −8 | 4 |
| | 17 | 6 | −2 | 0 | 0 |
| A | 20 | 10 | 11 | 4 | 16 |
| | 38 | 24 | 17 | 0 | 24 |
| B | 10 | 13 | 22 | 15 | 19 |
| | 6 | 15 | 18 | 25 | 33 |
| C | 19 | 9 | 8 | 9 | 8 |
| | 27 | 7 | 8 | 13 | 2 |
| D | 29 | 15 | 10 | −1 | 15 |
| | 23 | 25 | 12 | 11 | 19 |

分散分析表

| 変動要因 | 変動 | 自由度 | 分散 | 観測された分散比 | P−値 | F 境界値 |
|---|---|---|---|---|---|---|
| 標本 | 1522.08 | 4 | 380.52 | 11.74444444 | 1.66E−05 | 2.758711 |
| 列 | 866.88 | 4 | 216.72 | 6.688888889 | 0.000837 | 2.758711 |
| 交互作用 | 1465.92 | 16 | 91.62 | 2.827777778 | 0.009702 | 2.069086 |
| 繰り返し誤: | 810 | 25 | 32.4 | | | |
| | | | | | | |
| 合計 | 4664.88 | 49 | | | | |

標本はここでは薬を表し，P値が 0.01 より小さいので，薬は 1%の有意水準で有意であり，薬の効果が認められた．列はここでは餌の違いを表し，P値が 0.01 より小さいので，餌の違いによる効果も 1 %の有意水準で認められた．交互作用も 1%の有意水準で認められた.

このように繰り返しのある二元配置では，分散分析によって，薬，餌のそれぞれの主効果，薬と餌の間の交互作用，誤差変動（偶然誤差など）に分けて解析でき，情報量が増える．さらに誤差変動を小さくできるから精度も高くなる．できる限り，このような分散分析のできる形で実験を計画するのが望ましい.

④　一元配置と回帰分析

分散分析で要因が変量であるときは，分散分析の後に回帰分析する方がよい．回帰分析の方がより鋭敏であること，回帰直線を求めることによって，実験で使わなかった変量の場合の反応も推定できるからである.

水稲を 5 段階の施肥量で育てて，収量を調べた（4 反復の実験）.

5段階の窒素施肥量(g/m$^2$)で水稲の収量実験(kg/10a)をした.

| 処理 施肥量 | 反復 | | | | 平均 |
|---|---|---|---|---|---|
| 20 | 725 | 650 | 682.5 | 662.5 | 680 |
| 15 | 680 | 642.5 | 670 | 677.5 | 667.5 |
| 10 | 710 | 652.5 | 590 | 637.5 | 647.5 |
| 5 | 545 | 487.5 | 565 | 532.5 | 532.5 |
| 0 | 515 | 430 | 392.5 | 502.5 | 460 |

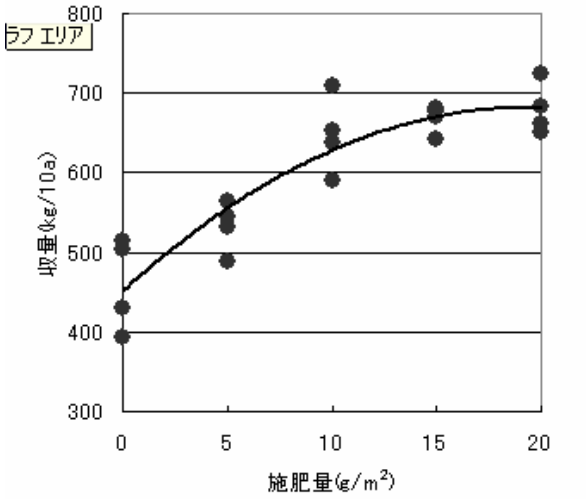
分散分析表

| 変動要因 | 変動 | 自由度 | 分散 | 観測された分散比 | P−値 | F 境界値 |
|---|---|---|---|---|---|---|
| グループ間 | 149350 | 4 | 37337.5 | 22.41370685 | 3.49E−06 | 4.893195 |
| グループ内 | 24987.5 | 15 | 1665.833 | | | |
| | | | | | | |
| 合計 | 174337.5 | 19 | | | | |

このようなエクセルの出力をふつうの分散分析表に直すと，以下のようになる．

| 要因 | 自由度 | 平方和 | 平均平方（分散） | 分散比（F 値） |
|---|---|---|---|---|
| 回帰 | 4 | 149350.0 | 37337.5 | 22.41*** |
| 残差 | 15 | 24987.5 | 1665.833 | |
| 全体 | 19 | 174337.5 | | |

量的因子の場合，実験で採用した水準（施肥量 20,15,10, 5,0g/m²）のうちのどれかが最適であるというわけではなく，実験で採用した水準と水準の間に最適な水準があるほうがむしろふつうである．右の図のように 2 次曲線で回帰できるならば，収量を最大にする施肥量は 19g/m² あたりにあることがわかる．




⑤　二元配置と回帰分析（乱塊法の場合）

④で扱った水稲の収量実験は実際には乱塊法で行ったものである．繰り返しのない二元配置の分散分析を行った結果は以下の通りであった．

| 処理 | ブロック | | | | | |
|---|---|---|---|---|---|---|
| 施肥量 | I | II | III | IV | 合計 | 平均 |
| 20 | 725 | 650 | 682.5 | 662.5 | 2720 | 680 |
| 15 | 680 | 642.5 | 670 | 677.5 | 2670 | 667.5 |
| 10 | 710 | 652.5 | 590 | 637.5 | 2590 | 647.5 |
| 5 | 545 | 487.5 | 565 | 532.5 | 2130 | 532.5 |
| 0 | 515 | 430 | 392.5 | 502.5 | 1840 | 460 |
| 合計 | 3175 | 2862.5 | 2900 | 3012.5 | 11950 | |
| 平均 | 635 | 572.5 | 580 | 602.5 | | 597.5 |

分散分析表

| 変動要因 | 変動 | 自由度 | 分散 | 観測された分散比 | P−値 | F 境界値 |
|---|---|---|---|---|---|---|
| 行 | 149350 | 4 | 37337.5 | 34.00759013 | 1.85E−06 | 5.411948 |
| 列 | 11812.5 | 3 | 3937.5 | 3.586337761 | 0.046571 | 5.952529 |
| 誤差 | 13175 | 12 | 1097.916667 | | | |
| | | | | | | |
| 合計 | 174337.5 | 19 | | | | |

以上のようなエクセルの出力をふつうの分散分析表に直すと，以下のようになる．

| 要因 | 自由度 | 平方和 | 平均平方（分散） | 分散比（F 値） |
|---|---|---|---|---|
| 処理 | 4 | 149350.0 | 37337.5 | 34.01** |
| ブロック | 3 | 11812.5 | 3937.5 | 3.59* |
| 誤差 | 12 | 13175.0 | 1097.9 | |
| 全体 | 19 | 174337.5 | | |

分散分析の結果から，処理間に 1%の有意水準で有意差が認められた．さらにブロック間にも有意差があったので，ブロックにすることで実験の誤差が減少し，実験の精度が向上したことがわかった．この場合，施肥量は量的因子であるから，施肥量に対して，収量がどのように反応したかを回帰分析できる．

この実験では乱塊法によって全体の変動の一部をブロックによる変動として除去できるので，回帰の分散分析は少し複雑になる．さらに直線でなく，２次式を回帰式として当てはめようとしてので，実際には以下のような分散分析表が得られた．

| 要因 | 自由度 | 平方和 | 平均平方（分散） | 分散比（F 値） |
|---|---|---|---|---|
| 処理 | 4 | 149350.0 | 37337.5 | 34.01** |
| 1 次回帰 | 1 | 132220.1 | 132220.1 | 120.43** |
| 2 次回帰 | 1 | 13207.7 | 13207.7 | 12.03** |
| 残差 | 2 | 3908.6 | 1954.3 | 1.78 |
| 誤差 | 12 | 13175.0 | 1097.9 | |

処理のうち，１次回帰で説明できる部分（平方和）がほとんどであるが，２次回帰も分散比が 1%水準で有意であったので，２次回帰とした．したがって，グラフのように２次曲線が得られ，施肥量 19kg のあたりで収量が最大となった．残差（処理による変動のうち，１次回帰および２次回帰では説明できない部分）は誤差と同じくらい小さくなったので，１次回帰および２次回帰で処理の変動を説明でき，これ以上の項を加える必要はない．

繰り返しのある回帰分析の場合，残差を計算できる．残差とは処理による変動のうち回帰では説明できない部分である．繰り返しのない場合，誤差（処理によらない変動であり，偶然誤差などを含む）と残差は分離できないので，分散分析表の誤差は誤差＋残差である．繰り返しがあると残差が計算でき，残差が大きいとその回帰式は当てはまりが悪いということになる．

したがって，

総変動＝処理による変動＋誤差変動（処理によらない変動）

＝回帰で説明できる変動＋残差（回帰で説明できない処理による変動）＋誤差変動

回帰分析できるデータでは分散分析の後，回帰が有意であれば処理による効果はどの水準でもみられると判断してよい．すなわち回帰式の傾きから，独立変数をわずかでも変化すれば，その傾きだけ従属変数が変化すると考えることができる．したがって，回帰分析の場合，水準間の差を多重検定する必要はない．先ほどの水稲の収量の例では 0 と 5kg の施肥間に多重検定で有意な差があるかを検定する必要はない．施肥量を変えれば収量が変わることが回帰分析からわかっているからである．さらに回帰分析では回帰式を用いれば 0 と 5kg の間の任意の施肥量で収量がいくらになるかを推定できる．

**Ｅ．定期試験について**

日時　2 月 1 日（火）　午後 3 時 30 分から 5 時まで（90 分間）
場所　マルチメディア演習室**1**（２ではありません）
すべて持ち込み可．パソコンは必ず持ってくること．ただし通信機能の使用は不可．
フロッピーで答案を出します.（フロッピーおよびＵＳＢ接続のフロッピーディスクドライブをこちらで用意します.）
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 31 日（月）午後５時