

実験計画と分散分析ANOVA

－平均値の差の検定－

村瀬 洋一

1. 分散分析とは何か

1.1. 分析の目的と具体例

◆目的 — 説明変数（独立変数）Xを複数設定し、被説明変数（従属変数）Yとの関連が強いのが、どの変数なのかを解明する。

ただし、Yは連続変数、Xは離散変数（カテゴリー）。

重回帰分析 — XもYも連続変数（量的）。両者は似ているがこの点が異なる。

分析法の考え方 — 他の説明変数Xの影響を統計的に取り除いた上でも（統計的統制をした上でも）、ある説明変数Xと被説明変数Yが関連しているか（Yに対するXの効果があるか、つまりX内のグループ間で、Yの平均値に有意な差があるか）を解明する。

具体的には、2つ以上の平均値の差の検定を行う。

- ・帰無仮説 各グループ（級、組、群、カテゴリー）の平均値は全て等しい。
- ・対立仮説 各平均値の間に差がある。

◆具体例 — ある調査データの意識に関する項目や所得等、連続変数として扱えるものをYとし、性別、職業、学歴カテゴリー、年齢カテゴリーなどをXとして分散分析を行う。例えば、テストの点数Yが、性別や、学歴3グループ（中、高、大）の間で異なるか分析。

例えば、ホワイトカラー、ブルーカラー、農業の3カテゴリー（職業X）で、Yの値（点数や身長や収入、ある意識の回答平均など）が異なるかについて分析する。つまり、3つの平均値の間に違いがあるかを分析する。違いがあれば、XはYと関連があるといえる。この例では、Xの数は1つで、X内のカテゴリーが3つある。

分散分析は実験と深く結びついた分析法だが、調査データについて使用することもある。実験データは、データの人数が少ないため、関連の強さは検討せず、差が有意かどうかのみを検討することが多い。また交互作用について検討する。ただし、実験でも調査データでも、交互作用を検討するならば、Xの数は3つ以内くらいが限界である。重回帰分析では、Xが10個以上あることも珍しくないが、この点が異なる。

なおt検定は2つの平均値の差の検定。分散分析のF検定は3つ以上も可。両者は必ず同じ結果になる。

1.2. 実験計画法の考え方

実験：現象間の因果関係を解明するため、原因と思われる現象に意図的に操作を加え、結果と思われる現象の変化の有無を観察する手続き。

例：子どもの攻撃性の実験　－　子供を2グループに分ける

暴力場面の多い映画を見せる群　→　攻撃行動が出た　（実験群）

暴力場面のない映画を見せる群　→　攻撃行動はなかった（統制群）

∴　子どもの攻撃性（結果）を育てる原因（要因）は暴力場面である

操作された要因　－　説明変数　X　（要因、独立変数：independent variable）

観察された要因　－　被説明変数　Y　（従属変数：dependent variable）

実験では、刺激（暴力場面）と反応の関連を見ることになる。サンプルを2つ以上のグループに分けるのが実験の特徴。したがって、2以上の平均値について、差の検定を行う。実験は、実験群の他に必ず統制群を作る。片方しかないとは実験にならない。この例ではXは1つだけ（暴力的映画）なので簡単だが、Xが複数あると、複雑な実験計画となる。

被説明変数Yにおいて観察された変化が、Xによって確実に引き起こされたと考えるためには、X以外のすべての条件が等しいことが必要である。しかし、実際には無数の条件の違いが混入してくる恐れがある。例えば、たまたま片方のグループに、活発な子供がたくさんいた、など事情である。その中でYに影響を及ぼすと仮定できるものを剰余変数（extraneous variable）と呼ぶ。剰余変数の有力候補は以下である。

被験者変数．．．性、年齢、性格、能力や知能、本人のやる気や動機づけ

それ以外．．．実験環境、課題条件、実験者のふるまい、教示

剰余変数を無効にする方法には、次の4種類がある。

（例　X：新しい教授法、Y：学習成績、剰余変数：被験者の知能）

1 均衡化　より多数の被験者を、実験群と統制群に無作為に割り付ける

2 相殺化　同じ知能水準の被験者が実験群と統制群で同数ずつになるよう配置する

3 恒常化　一定の知能水準の被験者だけを集めて実験する

4 説明変数化

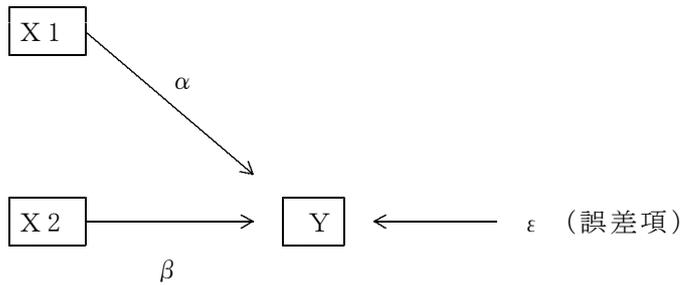
被験者を高知能群・（中知能群）・低知能群に分け、もう一つの説明変数にする

現実には、心理学実験の多くは、大学生のみからデータを取る。そこで、全員の年齢や学歴や能力は、ほぼ等しい（すでに恒常化した）と考えた上で実験を行うことが多い。

1.3. 分散分析の考え方

実験を行なった結果、被説明変数に違いが見られる場合、その違いが偶然によって生じたのか、それとも実験処理によって生じたと考えた方がよいのかについて検討するため、統計的な推定を行なう。男女など2つの平均値の比較にはt検定を用いるが、処理水準（level；X内のカテゴリー数）が3つ以上の場合や、Xが2つ以上存在する場合の平均値の差の検定には、分散分析（Analysis of Variance:ANOVA）を用いる（なお、説明変数Xが1つで、カテゴリー数2ならば、1元配置2水準の分析と呼ぶ）。

2元配置分散分析の基本モデル（グラフによる表現）



簡略化のため、Xが1つの場合を数式で表すと次のようになる。

分散分析の基本モデル（完全無作為法、水準数がm個で要因数1の場合）。

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \quad (i=1, \dots, n; j=1, \dots, m) \dots \dots \dots (1)$$

- y_{ij} : j番目の水準のi番目の個体の値（各個人の測定値）
- μ : 母集団の平均
- α_j : 母集団におけるj番目の水準の効果（グループに属する効果）
- ε_{ij} : j番目の水準のi番目の値に関する誤差

$\mu = \bar{y}$ （全体平均）、 $\alpha_j = \bar{y}_j - \bar{y}$ （全体平均と各水準の平均値の距離）と置き換えると、

$$y_{ij} - \bar{y} = (\bar{y}_j - \bar{y}) + (y_{ij} - \bar{y}_j) \dots \dots \dots (2)$$

（2）式の左辺は、各個人の値と平均値との距離（ばらつき）である。

両辺を2乗してiとjについて総和して整理すると、

$$\sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{y})^2 = n \sum_{j=1}^m (\bar{y}_j - \bar{y})^2 + \sum_{j=1}^m \sum_{i=1}^n (y_{ij} - \bar{y}_j)^2 \dots (3)$$

すなわち（3）式は、**SS total = SS between + SS within** となる。

全体平方和が級間平方和と級内平方和（残差平方和）に分解されることがわかる。

- SS total（全体平方和）：全体平均の回りの個々の測定値のばらつき
- SS between（級間平方和）：全体平均と各水準の平均値との距離（ばらつき）
- SS within（級内平方和）：各水準の平均値の回りの個々の測定値のばらつき
- SS : Sum of Square 二乗和

また偏差平方和（SS）をその自由度で割った商が平均平方（MS）である。

この平均平方の期待値は、各々の自由度で χ^2 分布することが知られている。

分散分析の具体的目的は、あるXの効果が有意であるか、すなわち各水準の平均値の差が「偶然のものではない」と見なしてよいか、の検定を行なうことである。これはすなわち、級間平均平方（分散）が級内平均平方（分散）より十分に大きいかどうかについて、検討しているのである（各水準の分散に関する図は村瀬他(2007:101)を参照）。

1.4. F 値（F 比）の考え方

$$F = \frac{MS_{\text{BETWEEN}}}{MS_{\text{WITHIN}}} = \frac{\text{級間平均平方}}{\text{級内平均平方}} = \frac{\text{モデルによって説明できる分散}}{\text{モデルによって説明できない分散}} \quad \dots (4)$$

この F 値は、 χ^2 分布する 2 つの変数の比であり、自由度 (m-1, n-m) の F 分布する。F 分布表で、F 値の出現する確率 p が、帰無仮説（各平均値に差はないという仮説、つまり各平均値の差は偶然という仮説）が採用される確率（危険率、有意水準）となる。F は 20 世紀最大の統計学者と言われるフィッシャーの名前から来ている。

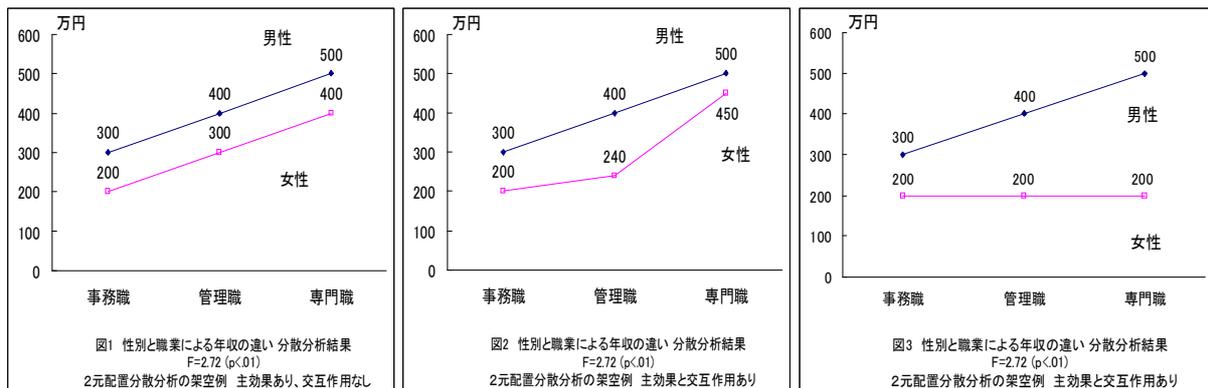
1.5. 実験計画の様々なデザイン

実験計画における説明変数の設定の仕方には、処理条件ごとに異なった被験者群を割り付ける完全無作為法 (completely randomized design) とすべての同一の被験者群にすべての処理条件での測定を実施する乱塊法 (randomized block design) がある。乱塊法は完全無作為法よりも被験者数を節約できる利点がある反面、先に条件での測定が後の条件での測定に影響を与える可能性があるため、その影響を相殺化などの操作で取り除くことができると考えられる場合以外には用いられない。またこれらの両方の種類の説明変数を用いた混合法 (mixed design) と呼ばれる計画もある。

1.6. 交互作用

複数の説明変数を組み合わせた要因計画 (factorial design) では、それぞれの要因 X の単独の影響（主効果）を検出できるほか、一方の要因の効果が他方の要因の水準によって異なるかどうか（交互作用）についても検討することが可能である。

以下の図は、要因 X1（性別）と要因 X2（職種）の組合せごとに、Y の平均値をグラフ化し、主効果と交互作用の関係を模式的に示したものである。2 つの線が平行なときは交互作用はない。重回帰分析と違って、分散分析は交互作用を容易に検討できる。



1.7. 多重比較

分散分析において（交互作用が有意でなく）有意な主効果があったということは、「その要因の各水準の平均がすべて等しい」という帰無仮説が棄却されたことを意味するにす

ぎない。説明変数が「10 m, 20 m, 30 m」のような比例尺度であったとしても、主効果が有意であることは、説明変数と被説明変数の間に線形関係があることを意味しない。両変数の関係がU字形または逆U字型の場合にも主効果は有意になる。さらに、Xの水準が3つ以上ある場合、どの水準とどの水準の間に統計的に有意な違いがあるかを知ることが必要になる。ここで通常のt検定を繰り返して実施すると、第1種の誤差に関して「甘い」検定になってしまう（検定の多重性）。

これを避けるため、全体としての有意水準を一定に保つための多重比較の方法が、各種考案されている。統計学者Tukeyの学生化した範囲の検定（テューキー法）、Scheffeの対比による方法、Bonferroniのt検定などがよく用いられる。

なおSPSSでは、「その後の検定」というボタンが多重比較のことである。

1.8. 分散分析の適用条件

分散分析では1)標本が正規分布にしたがう母集団から抽出されたものであること（正規性）、2)各水準にはいる標本が独立であること（独立性）、3)各水準にはいる標本の分散が等しいこと（等分散性）、の3つが前提とされる。正規性と等分散性については、完全に満足されなくても分析結果にたいして影響しない（頑健性がある）ことが知られているが、あまりにはずれる場合には分散分析は行うべきでない。測定の性質上の理由で、正規性や等分散性が崩れる場合には、測定値を適当に変換することにより、条件を満たすことができる場合がある。テストなど一定数の項目中の正答率を角変換したり、反応時間を対数変換することはよく行なわれる。また、標本数があまりに少ない場合や順序尺度である場合は分散分析を用いずに、より制約の緩いノンパラメトリック検定を用いる方がよい。

1.9. 非実験データの分析

上記のように、分散分析は元来、実験による因果関係の推論のために開発された統計技法であるが、調査のような非実験データの分析にも非常に有益である。ある変数Y（例えば所得）についてのデータを、別の変数X（例えば職業）など一定の基準で分割してグループをつくり、グループを説明変数Xとして平均値の差の検定を行う。これは個々人のYの値がばらつく（人によって違う）内、グループに属することによる違いとして説明できる分が統計的に有意かどうかを、検討することになる。非実験データの場合、説明変数Xは操作されたものではなく、剰余変数の統制が困難であるので、XとYの間に因果的な関係を推論しようとする際は慎重にすべきである。ただし実験データだからといって、関連が単なる相関でなく因果であると、断定することはできない。他の要因Xによる見せかけの相関である可能性も常にある。ただ調査データの方が、多くの要因を含むことが多い。

分散分析は本来、実験計画法と密接に結びついた分析法である。しかし社会調査データでも、カテゴリー別の平均に差があるかを分析するために、分散分析をよく用いる。調査データなど、比較的人数が多い大標本では、差や関連の大きさについて、検討することを目的とする。実験など小標本データは、差の大きさや関連の強さは検討できないので、差が有意かどうかを解明することに目的をしばって、分析することが多い。

2. SPSSによる分析

2.1. SPSSの操作

分散分析は、計算方法が比較的単純であるため電卓や表計算ソフトなどで行うことが可能である。統計パッケージのSASやSPSSには分散分析を行うANOVAという名前のプログラムが用意されている。SPSSでは、説明変数が1つだけの分散分析の場合、画面上の「分析」をクリックして、平均値の比較を選び、1元配置分散分析（シンタックスはoneway）を選ぶ。

あるいは、以下のようにシンタックスを書いても良い。太字の所に用いる変数を入れる。この例では、教育年数eduが被説明変数、年代nendaiが説明変数である。

重回帰分析の場合、説明変数は量的変数なので、年齢はなるべく細かい方が良い。しかし分散分析では、説明変数はカテゴリ変数なので、あまり細かい変数だと良くない。この例では、説明変数は年代という20代から60代まで5段階の変数を用いている。

シンタックス例 1元配置分散分析

```
ONEWAY  edu BY nendai          (←ONE だけでも動く   この例ではEDUがY)
/POSTHOC = TUKEY ALPHA(.05)
/STATISTICS DESCRIPTIVES      ← なくてもよい。平均値などを出す命令文
/PLOT MEANS .                  ← なくてもよい。平均値折線グラフを出す命令文
```

解説 1行目ONEWAY のあとに Y by X の順にモデルで使う変数を書く
/POSTHOC = TUKEY ALPHA(.05) 多重比較法にテューキー法を用い有意水準を5%とする
/PLOT MEANS . 各カテゴリー別平均値の図を出す

出力として、F値や有意水準、各年代グループの平均値等が出る。plot行を書くと折れ線グラフも出るので分かりやすい。また多重比較の表には、平均値間で有意な差があるところに*マークがついている。

2元配置以上の時は、「分析」をクリックして一般線形モデルを選び、「1変量」を選ぶ（シンタックスはunianova）。「多変量」は複数の従属変数を1度に分析するが、あまり使わない。

被説明変数として使いたいものを1つ選び、自分の好きなモデルを作ればよい。説明変数Xは普通、質的変数なので、固定因子のボックスに入れる（それ以外のボックスはあまり使わない）。共変量に入れるのは量的変数である。「モデル」ボタンを押して、分析に用いる説明変数を選ぶ（モデルは「すべての因子による」を選び、とくに指定しなくてもよい）。平方和はタイプⅢを選ぶことが一般的である。平方和の計算法に何種類かあるのだが、タイプⅢがよく使われる。交互作用は、まずは2次までを入れれば良い。

「オプション」ボタンを押して、記述統計を出すと、各カテゴリーごとの平均値の数字が出る。また「作図」ボタンを押して、横軸と線の定義変数を指定すると、各カテゴリーごとの平均値のグラフが出る。多重比較を行いたい時は「その後の検定」ボタンを押して変数を選ぶ。分析目的に応じてTukeyやScheffeなどを選ぶとよい。

あるいは、以下のようにシンタックスを書いても良い。太字の所に用いる変数を入れる。この例では、教育年数eduが被説明変数、年代nendaiと性別q1sexが説明変数である。

/DESIGN行にモデルを書けばよい。この例では、主効果として nendai q1sex 、交互作用

としてnendai*q1sex を入れたモデルにしている。この例では2次の交互作用のみを入れているが、説明変数が3つ以上ある時は、3次の交互作用を入れることもできる。多重比較をする時は/POSTHOC行を書く。以下の例ではSCHEFFE法で多重比較をしている。

シンタックス例 2元配置分散分析

例1 簡略な例

```
UNIANOVA  EDU BY  nendai q1sex      (←UNI だけでも動く この例ではEDUがY)
/DESIGN =  nendai q1sex  nendai*q1sex  ←自分の作ったモデルを書く
/POSTHOC =  nendai ( SCHEFFE )        ←多重比較をする場合に手法を書く
/PLOT = PROFILE( nendai*q1sex )      ←平均値の図を出す場合に書く
/PRINT = DESCRIPTIVE .               ←平均など記述統計量を出す場合に書く
```

例2 詳しい例

```
UNIANOVA  EDU BY  nendai q1sex      (←UNI だけでも動く この例ではEDUがY)
/METHOD =  SSTYPE(3)
/DESIGN =  nendai q1sex  nendai*q1sex  ←自分の作ったモデルを書く
/POSTHOC =  nendai ( SCHEFFE )        ←多重比較をする場合に手法を書く
/CRITERIA = ALPHA( .05)              ←多重比較の有意水準を書く (省略可)
/PLOT = PROFILE( nendai*q1sex )      ←平均値の図を出す場合に書く
/PRINT = DESCRIPTIVE .               ←記述統計量を出す場合に書く
```

以下が出力例。「被験者間効果の検定」という表が、いわゆる分散分析表。この例では、1行目がモデル全体のF値であり、19.776で有意であることが示されている。主効果はnendai 38.383、q1sex 39.053であり、どちらも有意。交互作用のF値は有意ではない。

被験者間効果の検定(分散分析表の例)

従属変数 本人学歴	タイプ III 平方和	自由度	平均平方	F 値	有意確率
修正モデル	1416.723	9	157.414	19.776	0.000
切片	203105.226	1	203105.226	25515.805	0.000
NENDAI	1222.108	4	305.527	38.383	0.000
Q1SEX	310.858	1	310.858	39.053	0.000
NENDAI * Q1SEX	54.265	4	13.566	1.704	0.147
誤差	8198.776	1030	7.960		
総和	226799.000	1040			
修正総和	9615.499	1039			

a R2 乗 = .147 (調整済みR2乗 = .140)

「被験者間因子」という表が別に出るが、これは各グループの人数Nが出ているだけである。記述統計量はグループごとに平均値等が出るだけ。シンタックスでplotオプションを書くと、平均値の折れ線グラフが出るので分かりやすい。また「その後の検定」という表は多重比較の結果である。平均値間に有意な差があるところに*マークがついている。

2.2. 結果のまとめ方

分散分析の結果は変動因、平方和 (SS)、自由度 (df)、平均平方 (MS)、F 値、有意水準 (p) などが一覧可能な、上記のような分散分析表という形式で表示する。ただし字数に制限のある雑誌論文などでは「F(2, 37)=9.67 (p<.01)」と省略して報告すること

もある（Fの後の括弧内の2つの数字は要因の自由度と誤差の自由度を示す）。

調査結果について、カテゴリーごとの平均値を提示する場合、グラフにすると分かりやすい。上記の図のように平均値を折線グラフにすると良い。図タイトル下の部分に注として、F値や自由度DF、有意水準pを書くこと。

3. 分析時の注意点

3.1. 分析の前に必ず欠損値処理をすること

多くの場合、欠損値は9か99。SPSSの場合、missing valuesコマンドを用いる。

回答が2桁の場合、欠損値99である。まず単純集計をとって確認するとよい。

3.2. 分析の前に変数の向きを必要に応じて逆転し、わかりやすく設定する

分析を行う前に、原則として、すべての変数を、数字が大きいほど肯定になるように直すこと。数字が小さいほど肯定となる変数が混ざっていると、とても分かりにくい。

3.3. 用いる変数について

用いる変数は、Yは連続変数（量的変数）、Xは離散変数（カテゴリー変数）であることに注意。各カテゴリー内の人数が少ない場合はカテゴリー合併を行うとよい。あるいは、人数が少なく異質すぎるカテゴリーは、はずれ値だと考えて分析から削除しても良い。

3.4. モデルの考え方とデータ人数

重回帰分析と分散分析は、数学的にはほとんど同じモデルである。しかし、重回帰分析は線形モデル（直線）を前提としているが、分散分析は、そのような前提にもとづいた分析法ではない。

重回帰分析は、関連の大きさを見るため、普通、400人以上の大きなデータでないと使わない。それに対して分散分析は、もともと少人数データの有意差検定のために作られたものである。有意かどうかを見るためには、数十人のデータでも良い。結局、分散分析とは、グループ間の平均値の差を検定しているだけである。平均値に差があるかどうか（有意かどうか）を分析しているだけで、関連の大きさについては、通常、考慮しない。大きさを示すイータ二乗（相関比）という係数を出すことができるが、あまり使われない。

重回帰分析や因子分析は、経験的には、最低でも400人以上のデータが必要である。人数が少ない場合は、分析結果は不安定になるので信用できない。少人数のデータで無理に多変量解析をやるべきではなく、人数が少ないならば、クロス集計等で十分である。

4. 課題

調査項目の中から、自分が分析に用いたい変数を決める（被説明変数Yは、連続変数を1つ決め、Xはカテゴリカル変数を2つ以上決める）。そしてプログラムを動かし分析を行う。何を説明したいか（Y）を決め、それを説明するのにふさわしい変数Xが何かを考え、自由にモデルを作ること。何種類かのモデルを自分で考え、試行錯誤すると良い。まずは、2次の交互作用のみを入れたモデルで分析すると良いだろう。

分析時には、上記の注意点に気をつけること。分析結果が出たら、F値や有意水準 p を見て、上記の図のように、各グループの平均値を、折れ線グラフにして結果をまとめる。Yの値を縦軸としてグラフにすればよい。SPSS出力でいくつかの平均値を出し、エクセルでグラフを作ること。

そして、各Xの主効果が有意か、交互作用は有意か、モデル全体の説明力はどうか等についてグラフ下を書く。また、結果の解釈や自分なりの考察を書くこと。解釈として自分独自の意見を書くことが重要である。

重回帰分析の課題と同様、データを男女に分割した後で、分析してもよい。

5. 参考文献

ボーンシュテット・ノーキ. 1990. 『社会統計学 —社会調査のためのデータ分析入門』. ハーベスト社.

市川伸一・大橋靖雄・岸本淳司・浜田知久馬. 1993. 『SASによるデータ解析入門 第2版』東京大学出版会.

石村貞夫. 1992. 『分散分析のはなし』東京図書.

石村貞夫. 2002. 『SPSSによる分散分析と多重比較の手順 第2版』東京図書.

石村貞夫. 2004. 『SPSSによる統計処理の手順 第4版』東京図書

海保博之編. 1985. 『心理・教育データの解析法10講 基礎編』福村出版.

岸学. 2005. 『SPSSによるやさしい統計学』オーム社.

★分散分析について解説が詳しい

小牧純爾. 1995. 『データ分析法要説 —分散分析法を中心に』ナカニシヤ出版.

三宅一郎他編. 1991. 『新版SPSS X第3巻 解析編2』東洋経済新報社.

村瀬洋一他編. 2007. 『SPSSによる多変量解析』オーム社.

★各カテゴリーごとの分散（ばらつき）についての図はこれを参照

小野寺孝義,・山本嘉一郎編. 2004. 『SPSS事典』ナカニシヤ出版.

高橋行雄・大橋靖雄・芳賀敏郎. 1989. 『SASによる実験データの解析』東京大学出版会

豊田秀樹. 1994. 『違いを見抜く統計学 —実験計画と分散分析入門』講談社ブルーバックス.

その他、SPSS操作法やシNTAX解説については、重回帰分析の資料の参考文献や、村瀬他(2007)を読むこと。