

研究法や分析法に関するメモ

奥村太一

平成 29 年 9 月 27 日

1 LOCF 解析と ITT 分析

1.1 LOCF 解析

LOCF 解析 (last observation carried forward analysis) は、臨床試験における欠測データの処理方法の 1 つであり、欠測データを直近の先行観測値で補完して解析する方法である。以下、外部サイト¹より引用。

LOCF 解析は、欠測データを直近の先行観測値で補完して解析する方法である。たとえば、投与開始時から 1 週間ごとに 4 週目まで血圧を測定する場合を考える。4 週目の測定値がない被験者は 3 週目の測定値があれば、4 週目を 3 週目の測定値で代用する。もし、3 週目から 4 週目の間で、血圧の経時的な推移が、ほぼ一定であると仮定できればよいが、増加傾向や減少傾向が疑われる場合には問題がある。

なお、ベースラインデータで欠測データを補完する方法は **BOCF 解析** (baseline observation carried forward analysis) と呼ばれる。

臨床試験におけるガイドラインについては、厚生労働省による「臨床試験のための統計的原則」について (医薬審 第 1047 号, 平成 10 年 11 月 30 日) (1998)²を参照のこと。

1.2 ITT 分析

臨床試験においては、諸般の事由から当初割り当てた条件を途中で中断したり別の条件に変更したりすることが生じる。

例えば、新しい治療と標準的治療の効果を比較しようとする場合に、

群	治療
A	新しい治療 → そのまま新しい治療を継続
B	新しい治療 → 標準的治療に変更
C	標準的治療 → そのまま標準的治療を継続
D	標準的治療 → 新しい治療に変更

といったように当初の割り当て通りにいかない被験者が発生する場合がある。

ITT 分析 (intention-to-treat analysis) とは、対象者が実際に割り付けられた治療を完結したか、あるいは実際にはじめから受けたか どうかにかかわらず、当初割り付けた群にしたがって分析すること³である。つまり、

[A+B 群] と [C+D 群]

¹<http://statmedoo.publog.jp/archives/4354107.html>

²<https://www.pmda.go.jp/int-activities/int-harmony/ich/0031.html>

³<http://kamiyacho.org/ebm/ce304.html>

の比較を行う。

上記外部サイト⁴によれば、

対象者が最初に無作為に割り当てた治療方針に従わなかったのには何らかの理由があるためであり、intention-to-treat に従って分析するということは、治療そのものというよりは、そのマイナスの治療効果も含めて「治療方針」を比較しているものであり、この治療方針の効果は「実践的な治療効果」ともよばれています。医療・保健機関として実際に知りたいのは、その治療を実施する際のサービスそのものに対する実践的な治療効果であり、むしろ Intention-to-treat 分析を行うほうが実用的といえます。

とされる。

2 一対比較法や順位法のための IRT モデル

2.1 反応バイアス修正とサーストンの一対比較モデル

荘島 (編) 「計量パーソナリティ心理学」第 4 章

測定後の反応バイアス修正として

- 社会的望ましき尺度 (Edwards, 1957)
- 黙従傾向尺度 (Cauch & Keniston, 1960)

などが知られているが、事前にバイアスを除去する方法として

- 一対比較法
- 順位法

が有用だという報告がある (e.g. Cheung & Chan, 2002)。

これらの方法で得られたデータはイブサティブデータであり、項目間に独立性を仮定することができない。サーストンの一対比較モデルが刺激の順序付けを行うために利用されてきたが、

- 心理的特性の測定にも使用できる
- IRT モデルとみなすことができる

ことがわかり、サーストン型 IRT モデルの拡張が行われている (Meydeu-Olivares, 2001; Maydeu-Olivares & Brown, 2010)。

2.2 サーストンの比較判断の法則

渡部洋「人文・社会科学の統計学」P.312-

刺激の尺度 S_j, S_k について心理学的連続体を仮定し、それぞれに対する反応 (X_j, X_k) について

$$\begin{pmatrix} X_j \\ X_k \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_j \\ \mu_k \end{pmatrix}, \begin{pmatrix} \sigma_j^2 & \rho_{jk}\sigma_j\sigma_k \\ \rho_{kj}\sigma_k\sigma_j & \sigma_k^2 \end{pmatrix} \right)$$

とする。このとき、

$$D_{jk} = X_j - X_k \sim N(\mu_j - \mu_k, \sigma_{jk}^2)$$

⁴<http://kamiyacho.org/ebm/ce304.html>

となる。(ただし $\sigma_{jk}^2 = \sigma_j^2 + \sigma_k^2 - 2\rho_{jk}\sigma_j\sigma_k$)。

これを用いて、刺激 S_j を S_k よりも好む確率は、

$$\begin{aligned} p_{jk} &= Pr[D_{jk} > 0] \\ &= Pr\left[\frac{(X_j - X_k) - (\mu_j - \mu_k)}{\sigma_{jk}} > -\frac{\mu_j - \mu_k}{\sigma_{jk}}\right] \\ &= Pr\left[Z_{jk} > -\frac{\mu_j - \mu_k}{\sigma_{jk}}\right] \\ &= \Phi\left[\frac{\mu_j - \mu_k}{\sigma_{jk}}\right] \end{aligned}$$

となる。

よって、もし p_{jk} が得られていれば、

$$\frac{\mu_j - \mu_k}{\sigma_{jk}} = \Phi^{-1}[p_{jk}]$$

と求められる。

ここで、サーストンのケース **V** では、

$$\sigma_j^2 = \sigma_k^2 = \sigma^2, \rho_{jk} = 0$$

とおく。このとき、 $\sigma_{jk}^2 = \sqrt{2}\sigma$ となる。従って、 $\sqrt{2}\sigma$ を単位とした心理学的連続体の尺度を考えれば、

$$\mu_j - \mu_k = \Phi^{-1}[p_{jk}]$$

となる。

はてな？

「 $\sqrt{2}\sigma$ を単位とした心理学的連続体の尺度を考えれば」というのがいまいちよくわからない。

はてな？ここまで

ここで、一対比較データから比較判断の法則を利用して m 個の刺激に対する一次元尺度を構成することを考える。

1. まず、データから p_{jk} を求める⁵。
2. 次に、 $z_{jk} = \Phi^{-1}[p_{jk}]$ を求める⁶。
3. 原点を $\bar{\mu} = \frac{1}{m} \sum_{k=1}^m \mu_k = 0$ とする。
4. $\sum_{k=1}^m z_{jk} = \sum_{k=1}^m (\mu_j - \mu_k) = m\mu_j - \sum_{k=1}^m \mu_k = m\mu_j$ を利用して、 $(\mu_1, \mu_2, \dots, \mu_m)$ を求める。

ただし、以下の点に注意

- $p_{jk} \approx 0, 1$ のときは μ_j の値に信頼がおけない
- 尺度値の精度に関する情報が欠けている
- 一次元性の保証はない

⁵ N 人中の n 人が $S_j > S_k$ と判断した、などとして求めるのであろう。

⁶ $\sqrt{2}\sigma$ を単位とした心理学的連続体の尺度を考えれば、 $z_{jk} = \mu_j - \mu_k$ である。

2.3 サーストン型 IRT モデル

荘島 (編) 「計量パーソナリティ心理学」第 4 章

一対比較の場合、3 つの刺激 A, B, C に対して、ある被験者 i の選択が

1. $A > B$
2. $A > C$
3. $B < C$

であれば、

被験者	X_{AB}	X_{AC}	X_{BC}
i	1	1	0

とデータ化する。この場合、順序に整合性があるとは限らない。

同じ刺激について順位法を行った場合、 $A > C > B$ であれば、

被験者	X_{AB}	X_{AC}	X_{BC}
i	1	1	0

とデータ化する。

ここで、「A と B の比較で、A を選ぶ」ということをモデル化する。

- A を選ぶ傾向: p_A
- B を選ぶ傾向: p_B

とし、

$$X_{AB} = \begin{cases} 1 & (P_A - P_B \geq 0) \\ 0 & (P_A - P_B < 0) \end{cases}$$

となるとする。

このとき、

$$P_A = \mu_A + \beta_{A1}\theta_1 + \delta_A \quad (2.1)$$

$$P_B = \mu_B + \beta_{B2}\theta_2 + \delta_B \quad (2.2)$$

とする。ただし、 μ は P の平均、 β は項目特性、 θ は心理特性、 δ は反応の個人差とする⁷。

すると、反応の不整合を生む偶然誤差 e_{AB} を加えて⁸、

$$P_A - P_B = \underbrace{(\mu_A - \mu_B)}_{\alpha_{AB}} + (\beta_{A1}\theta_1 - \beta_{B2}\theta_2) + (\delta_A - \delta_B) + e_{AB}$$

となる。

⁷ここでは、各項目は単一の θ を測定しているという独立クラスターが仮定されているが、複数の特性を測定しているという仮定のもとでもモデル化することも可能。

⁸順位法だとなし。

すると、 $P_A - P_B \geq 0$ は、

$$\alpha_{AB} + (\beta_{A1}\theta_1 - \beta_{B2}\theta_2) + (\delta_A - \delta_B) + e_{AB} \geq 0 \quad (2.3)$$

$$\underbrace{(\delta_A - \delta_B) + e_{AB}}_{\xi_{AB}} \geq -\{\alpha_{AB} + (\beta_{A1}\theta_1 - \beta_{B2}\theta_2)\} \quad (2.4)$$

$$\frac{\xi_{AB}}{\sigma_{\xi_{AB}}} \geq -\frac{\alpha_{AB} + (\beta_{A1}\theta_1 - \beta_{B2}\theta_2)}{\sigma_{\xi_{AB}}} \quad (2.5)$$

$$Z_{AB} \geq -\underbrace{\frac{\alpha_{AB} + (\beta_{A1}\theta_1 - \beta_{B2}\theta_2)}{\sigma_{\xi_{AB}}}}_{\text{threshold}} \quad (2.6)$$

と標準正規分布に従う確率変数として書き下すことができる。

よって、

$$\Pr[X_{AB} = 1] = \Pr[P_A - P_B > 0] = \Phi \left[\frac{\alpha_{AB} + (\beta_{A1}\theta_1 - \beta_{B2}\theta_2)}{\sigma_{\xi_{AB}}} \right]$$

となる。すなわち、 $X_{AB} = 1$ となる確率（B より A が選ばれる確率）は、標準正規分布に従う確率変数が閾値を超える確率に等しいという正規累積モデルが成り立つ。

2.4 文献

- Cheung & Chan (2002)：イプサティブデータによる反応バイアスの低減
- Meydeu-Olivares (2001)：イプサティブデータの IRT
- Meydeu-Olivares & Brown (2010)：イプサティブデータの IRT
- 豊田「項目反応理論：理論編」：イプサティブデータの IRT

3 一般均衡的な効果の存在

伊藤公一朗 (2017). データ分析の力 光文社新書 P.259 -

小規模に実施した RCT の効果が、大規模に政策を実践した場合に当てはまらない場合がある。

例 テネシー州における、少人数学級の効果を検証した RCT → 学力アップ (1986)

→ カリフォルニア州で少人数学級の政策が開始 (1996) → とても小さな効果しか得られなかったと考えられる理由として、

- カリフォルニアの結果が信用できない
- 外的妥当性が低かった

に加えて、

- 学級増加 → 教員需要増加 → 教員の質低下

の可能性が考えられる。

つまり、大規模な実践では、小規模の RCT では生じなかった想定外の介入効果が現れることがあるということ。

4 R^2 と η^2

Hedges & Olkin (1985) P.101

R^2 分散トータルのうち回帰によって説明できる部分

- X は量的, 線形関係の評価
- 多値も可能: 水準数 > 2 なら非線形関係もありうるが, 相関の 2 乗には必ずしも反映されない。

η^2 離散変数 X 所与のときの Y の条件付き分散を σ_Y^2 から除いた部分

- X と Y の線形, 非線形関係両方を含む

しばしば, $R^2 = \eta^2$ として解釈しがち。

しかし, 線形回帰にもとづくものと, 非線形も含まれるものを単純に置き換えるべきではない。

例えば, Cohen (1988) による重回帰分析のための効果量 f^2 の解釈基準は R^2 に変換できるが, これらはあくまで線形回帰にもとづく。

よって, この解釈基準を η^2 に適用するのはまずい。

Hedges & Olkin (1985) P.101-102

- X が 2 水準: $R^2 = \eta^2$
- X が 3 水準以上: $R^2 \leq \eta^2$

R^2 と η^2 は分散説明率としては同じ意味であるが, 拾っている情報は少し異なる。解釈に注意。

5 てこ比: 主成分分析による多変量外れ値の検出

5.1 てこ比

多変量外れ値を検出するために, 主成分分析を利用しててこ比 *leverage* を各ケースについて計算することが有効⁹。

てこ比 全ての主成分に関する標準化された主成分得点の平方和を標本サイズで割ったもの。

変数の数を p 、標本サイズを N 、 j 番目 ($j = 1, \dots, p$) の標準化された主成分得点を $\tilde{f}^{(j)}$ とすると、

$$LEVERAGE = \frac{\tilde{f}^{(1)2} + \tilde{f}^{(2)2} + \dots + \tilde{f}^{(p)2}}{N} \quad (5.1)$$

である。

一般に、

$$0 \leq LEVERAGE \leq 1 \quad (5.2)$$

$$\overline{LEVERAGE} = \frac{p}{N} \quad (5.3)$$

である。

⁹柳井晴夫・緒方裕光 (2006). SPSS による統計データ解析 現代数学社 P.191

てこ比の値が大きいほど全体の傾向から外れており、平均の 2 倍を超えるケース

$$LEVERAGE > 2 \times \frac{p}{N} \quad (5.4)$$

もしくは値が 0.5 を上回るケース

$$LEVERAGE > 0.5 \quad (5.5)$$

を外れ値とすることが多い。

5.2 R による解析

主成分分析は、関数 `princomp()` で実行できる¹⁰。引数として、

- `x` : データフレーム
- `cor` : 相関係数行列を用いるときは `T` に指定（デフォルトは `F` であり、分散共分散行列を用いることになっている）
- `scores` : 主成分得点を算出する場合は `T`（デフォルト）

を指定する。

結果には、以下の内容が含まれる。

- `$loadings` : 主成分負荷量（固有ベクトル）
- `$scores` : 主成分得点（平均ゼロになるよう変換済み）

これを用いて、`data01` からてこ比を算出し、平均の 2 倍を超えるケースを除いたデータ `data01.omit` を作成するには、次のようにすれば良い。

```
> prin <- princomp(x=data01, cor=T) # 相関行列を用いた主成分分析
> std.prin <- scale(prin$scores) # 標準化された主成分得点を算出
> leverage <- apply(std.prin^2, 1, sum)/dim(std.prin)[1] # てこ比の算出
> is.out <- (leverage > 2*mean(leverage)); sum(is.out) # 外れ値とその数
> data01.omit <- data01[is.out!=T,] # 外れ値を除去したデータ
```

6 多変量分散分析

6.1 概要

「SAS による実験データの解析」第 14 章

- 複数の特性（従属変数）いずれにも同様の関心がある場合に、特性間の相関を考慮して仮説検証を行う¹¹。
- ただし、MANOVA は手続きが大変な割に結果の解釈がうまくできない。通常は特性ごとに ANOVA を適用する¹²。
- 経時変化データも、各時点の観測値を別々の特性と考えることで、MANOVA を適用できる¹³。

¹⁰金明哲 (2007). R によるデータサイエンス 森北出版株式会社 P.69

¹¹多くの従属変数について検定を繰り返すことで、多重性が生じる。

¹²スクリーニングとして利用し、MANOVA で有意差があれば特性ごとに ANOVA という順でもよいだろう。

¹³球面性仮定 (H-F 条件) が満たされていないと対応あり ANOVA は適用できないし、自由度調整によってこれに対処しようとするとなれば複雑な分析に進めない。

6.2 独立な2群の比較（従属変数が2つ以上）

図 6.1 は、薬剤の投与によってラットの肝臓量と体重がどう変化するか検証するために集められたデータ `dataliverweight.csv` である。

	A	B	C	D
1	個体	薬	肝臓量	体重
2	1	x	710	245
3	2	x	627	224
4	3	x	865	342
5	4	x	1064	403
6	5	x	736	214
7	6	x	940	325
8	7	x	935	384
9	8	x	822	276
10	9	x	603	168
11	10	x	897	313
12	11	o	1074	347
13	12	o	806	220
14	13	o	1013	282
15	14	o	893	227
16	15	o	798	185
17	16	o	861	231
18	17	o	1162	393
19	18	o	1060	320
20	19	o	741	148
21	20	o	947	238

図 6.1: 薬剤の投与がラットの肝臓量（mg）と体重（g）に及ぼす影響

図 6.2 は薬剤の有無別にデータをプロットしたものである。体重が同じであれば、薬ありの方が肝臓量は高い傾向にある。

6.2.1 Hotelling の T^2

特性 k について、 i 群の j 番目の観測値を x_{ijk} とする。このとき、通常の t 検定では、検定統計量

$$t_k = \frac{\bar{x}_{1k} - \bar{x}_{2k}}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_k^2}} \quad (6.1)$$

(6.2)

を用いる。ただし、

$$S_k^2 = \frac{1}{n_1 + n_2 - 2} \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ijk} - \bar{x}_{ik})^2 \quad (6.3)$$

であり、特性 k についてプールされた群内不偏分散を表す。

ここで、

$$t_k^2 = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1} (\bar{x}_{1k} - \bar{x}_{2k}) (S_k^2)^{-1} (\bar{x}_{1k} - \bar{x}_{2k}) \quad (6.4)$$

と表すこともできる。

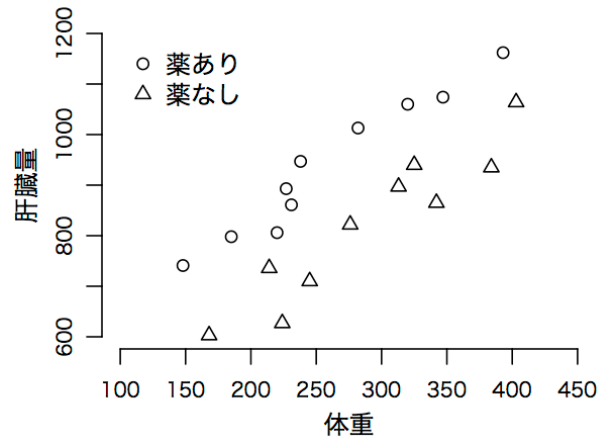


図 6.2: 群別プロット

例のように特性が 2 個の場合、

$$\bar{\mathbf{x}}_i = \begin{pmatrix} \bar{x}_{i1} \\ \bar{x}_{i2} \end{pmatrix}, \mathbf{V} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \quad (6.5)$$

と定義する。ただし、 v_{11} と v_{22} はそれぞれ特性 1 と特性 2 に関するプールされた群内不偏分散（つまり S_1^2 と S_2^2 ）、 $v_{12} = v_{21}$ は特性 1 と特性 2 に関するプールされた群内不偏共分散である。一般には、 $\mathbf{V} = \{v_{kk'}\}$ で

$$v_{kk'} = \frac{1}{n_1 + n_2 - 2} \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ijk} - \bar{x}_{ik})(x_{ijk'} - \bar{x}_{ik'}) \quad (6.6)$$

と書ける。

このとき、

$$T^2 = \left(\frac{1}{n_1} + \frac{1}{n_2} \right)^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{V}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad (6.7)$$

を Hotelling の T^2 統計量という。

もし帰無仮説 $H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ が真で、観測値ベクトルが共通の群内母分散共分散行列をもつ多変量正規分布に従うなら、

$$\frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)} T^2 \sim F(p, n_1 + n_2 - p - 1) \quad (6.8)$$

となり (p は特性の数)、 F 検定によってこの H_0 を検定できる。

R では、DescTools パッケージの関数 HotellingsT2Test() によって Hotelling の T^2 にもとづいた F 検定を行うことができる。

図 6.1 のデータ `dataliverweight.csv` で肝臓量と体重の平均が薬の有無に関わらず等しいという帰無仮説を検定するには、以下のスクリプトを実行すればよい。

```
> library("DescTools")
> head(data01,5)
  個体 薬 肝臓量 体重
1    1  x   710  245
2    2  x   627  224
```

```

3      3  x      865  342
4      4  x     1064  403
5      5  x      736  214
> HotellingsT2Test(cbind(肝臓量, 体重)~薬, data=data01)

```

Hotelling's two sample T2-test

```

data:  cbind(肝臓量, 体重) by 薬
T.2 = 41.086, df1 = 2, df2 = 17, p-value = 3.087e-07
alternative hypothesis: true location difference is not equal to c(0,0)

```

このアウトプットでは $T^2 = 41.086$ であるかのように表示されているが、本当は $T^2 = 87.00$ であり、T.2 の値は F 値である。結果は、 $F(2, 17) = 41.086, p = .000$ と書ける。

6.2.2 平方和積和行列の分割

特性が2つ以上の場合、ANOVA における平方和の分割を平方和積和行列の分割として以下のように拡張できる。

$$\mathbf{T} = \mathbf{H} + \mathbf{E} \quad (6.9)$$

ただし、 $\mathbf{T} = \{T_{kk'}\}$ は総平方和積和行列、

$$T_{kk'} = \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ijk} - \bar{x}_k)(x_{ijk'} - \bar{x}_{k'}) \quad (6.10)$$

$\mathbf{H} = \{H_{kk'}\}$ は群間平方和積和行列、

$$H_{kk'} = \sum_{i=1}^2 n_i (\bar{x}_{ik} - \bar{x}_k)(\bar{x}_{ik'} - \bar{x}_{k'}) \quad (6.11)$$

$\mathbf{E} = \{E_{kk'}\}$ は残差平方和積和行列

$$E_{kk'} = \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ijk} - \bar{x}_{ik})(x_{ijk'} - \bar{x}_{ik'}) \quad (6.12)$$

$$= (n_1 + n_2 - 2)v_{kk'} \quad (6.13)$$

である。

平方和積和行列の対角成分には特性 k に関する平方和が、非対角成分には特性 k と k' の積和（共分散 × 自由度）が格納されている。

このとき、

$$\mathbf{E}^{-1}\mathbf{H} \quad (6.14)$$

の固有値は、 $l = \min(p, q)$ 個得られる¹⁴。この固有値を $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l$ とすると、検定のための統計

¹⁴一元配置の場合、 $q =$ 群の数 $- 1$

量が 4 種類

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{H} + \mathbf{E}|} = \prod_i \frac{1}{1 + \lambda_i} \quad [\text{Wilks のラムダ}] \quad (6.15)$$

$$V = \text{tr}(\mathbf{H})(\mathbf{H} + \mathbf{E})^{-1} = \sum_i \frac{\lambda_i}{1 + \lambda_i} \quad [\text{Pillai のトレース}] \quad (6.16)$$

$$U = \text{tr}(\mathbf{E}^{-1}\mathbf{H}) = \sum_i \lambda_i \quad [\text{Hotelling-Lawley のトレース}] \quad (6.17)$$

$$\Theta = \lambda_1 \quad [\text{Roy の最大根}] \quad (6.18)$$

得られる。

例のように 2 特性 2 群比較だと、 $l = 1$ であり、

$$T^2 = (n_1 + n_2 - 2)\lambda_1 \quad (6.19)$$

が成り立つ。

Λ, V, U, Θ は $l = 1$ の時のみ等しい値を取り、かつ正確に H_0 の下で F 分布に従う。 $l = 2$ のときは Λ のみが正確。

6.2.3 正準変量

$\mathbf{E}^{-1}\mathbf{H}$ の固有ベクトルを用いて、もとのデータを l 個の正準変量に変換することができる。これらを C_1, C_2, \dots, C_l とすると、 C_1 は群内平方和が 1、群内積和が 0 となるように基準化されており、かつ群間平方和が最大値 λ_1 を取る。また、 C_2 は C_1 と無相関かつ群内平方和が 1、群内積和が 0 で、群間平方和が最大値 λ_2 を取る。以下同様。

群間差を視覚的に把握するため、特性数が 3 以上、群の数が 2 以上の場合、正準変量を 2 次元平面上にプロットすることがある。これを正準判別分析という。

R では、MANOVA による分析を行うための関数 `manova()` が装備されている。

図 6.1 のデータ `dataliverweight.csv` で肝臓量と体重の平均が薬の有無に関わらず等しいという帰無仮説を MANOVA によって検定するには、以下のスクリプトを実行すればよい。

```
> fit01 <- manova(cbind(肝臓量, 体重)~薬, data=data01)
> summary(fit01, test="Wilks")
              Df   Wilks approx F num Df den Df    Pr(>F)
薬              1 0.17142   41.086     2    17 3.087e-07 ***
Residuals 18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> summary(fit01, test="Pillai") # デフォルト
              Df  Pillai approx F num Df den Df    Pr(>F)
薬              1 0.82858   41.086     2    17 3.087e-07 ***
Residuals 18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> summary(fit01, test="Hotelling-Lawley")
```

```

          Df Hotelling-Lawley approx F num Df den Df      Pr(>F)
薬          1          4.8337   41.086      2     17 3.087e-07 ***
Residuals 18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> summary(fit01, test="Roy")
          Df      Roy approx F num Df den Df      Pr(>F)
薬          1 4.8337   41.086      2     17 3.087e-07 ***
Residuals 18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

この例は2群比較であるので、 $l = 1$ より4つの統計量いずれにもとづいても同じ結果が得られる。

上記の例ではMANOVAで群間差が有意であったので、次に特性ごとのANOVAを行うことになる。この場合、関数 `summary.aov()` に `manova()` の結果を渡せばよい。

```

> summary.aov(fit01)
Response 体重 :
          Df Sum Sq Mean Sq F value Pr(>F)
薬          1   4590   4590.4   0.7883 0.3863
Residuals   18 104813   5823.0

Response 肝臓量 :
          Df Sum Sq Mean Sq F value Pr(>F)
薬          1  66817   66817   3.2366 0.0888 .
Residuals   18 371599   20644

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

この場合、MANOVAでは有意差が得られていたものが、体重と肝臓量それぞれについて検討すると、有意差なしという結果になる。

上記の例で、平方和積和行列 \mathbf{H} , \mathbf{E} と $\mathbf{E}^{-1}\mathbf{H}$ の固有値は以下の通り。

```

> summary(fit01)$SS
$薬
      肝臓量      体重
肝臓量 66816.8 -17513.40
体重   -17513.4   4590.45

```

```

$Residuals
      肝臓量      体重
肝臓量 371599.4 189145.9
体重   189145.9 104813.3

> summary(fit01)$Eigenvalues
      [,1] [,2]
薬 4.833669    0

```

後藤 (1973, p.212) を参照して正準変量を算出すると、以下の通り。

```

> # E^(-1) %*% H の固有ベクトル
> H <- summary(fit01)$SS[[1]]
> E <- summary(fit01)$SS[[2]]
> inverseE.H <- solve(E) %*% H
> eigen.v <- eigen(inverseE.H)$vectors
>
> c.v <- cbind(scale(data01$肝臓量, scale=F), scale(data01$体重, scale=F)) %*%
+           eigen.v      # 正準変量
> colnames(c.v) <- c("c1", "c2")
> data02 <- cbind(data01[,1:2], c.v)
> head(data02, n=3)
  個体 薬      c1      c2
1    1  x -53.79489 -70.81384
2    2  x -74.67694 -112.17194
3    3  x -65.66752  62.31613

```

図 6.3 は薬剤の有無別に正準変量をプロットしたものである後藤 (1973, p.212) では各変数から全体平均を引いたものに固有ベクトルをかけることで正準変量を求めている。一方、「SAS による実験データの解析」図表 14.3 (p.275) ではこれとメトリックが異なっているし、そもそも SAS のアウトプットでは固有ベクトル (に何かを演算したもの) を用いて正準変量を算出している (図表 14.2, p.272)。この点については詳しく説明されていないのでよくわからない。

6.3 経時変化データ

6.3.1 Hotelling の T^2 と MANOVA による 2 群比較

- 分散共分散構造に H-F 条件のような仮定を必要としない
- 多重性を考慮した同時信頼区間を算出できる
- 標本数は時点数に比べある程度大きくないと検定力が低くなる

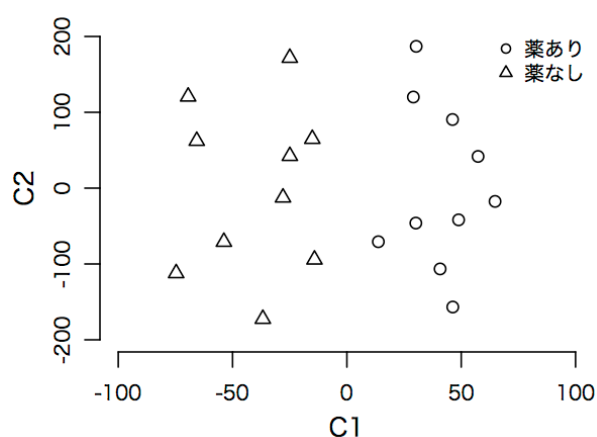


図 6.3: 群別正準変量プロット

- 欠測を含む個体は除去する必要がある
- H-F 条件が成り立っている場合は、通常の ANOVA の方が検定力が高い

図 6.4 は、降圧剤の投与量（10mg, 20mg）と血圧の経時変化（投与前、1 時間後、3 時間後）に関するデータ `databloodpressure.csv` を示したものである。

	A	B	C	D	E
1	被験者	降圧剤	投与前	一時間後	三時間後
2	1	10mg	119	113	114
3	2	10mg	110	115	112
4	3	10mg	123	126	111
5	4	10mg	130	127	100
6	5	10mg	121	115	124
7	6	10mg	135	125	115
8	7	20mg	125	109	113
9	8	20mg	108	98	104
10	9	20mg	133	118	117
11	10	20mg	125	114	110
12	11	20mg	113	107	106
13	12	20mg	126	118	110

図 6.4: 降圧剤の投与量と血圧の経時変化

図 6.5 は、このデータを箱ひげ図として表したものである。投与後時間が経過するごとに血圧が下がっていくこと、また 20mg 投与群の方が 10mg 投与群より全体的に血圧の低いことがわかる。低下の程度は 20mg 投与群の方が大きいようにも見えるが、判然とはしない。

今、Hotelling の T^2 と MANOVA によって、10mg 群と 20mg 群で血圧の母平均ベクトルが等しいという帰無仮説を検定する。すると、 $F(3, 8) = 5.5624, p = .02334$ となってこの帰無仮説は棄却される¹⁵。

ただし、母平均ベクトルが 2 群で等しくないという結果だけでは、「血圧の高さそのものが異なる」のか「血圧の下がり方が異なる」のかわからない。

R のスクリプトと実行結果は以下の通り。

¹⁵この場合も 2 群比較であるから $l = 1$ であり、いずれの方法でも結果は同じになる。

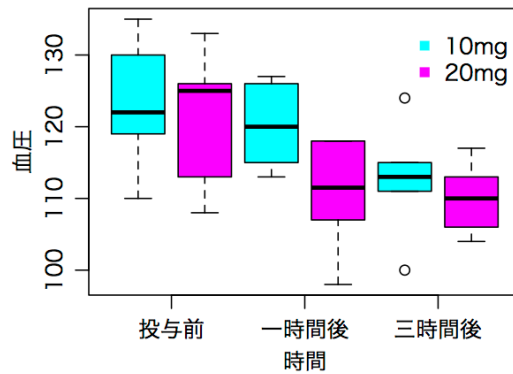


図 6.5: 箱ひげ図

```
> HotellingsT2Test(cbind(投与前, 一時間後, 三時間後)~降圧剤, data=data01)
```

Hotelling's two sample T2-test

data: cbind(投与前, 一時間後, 三時間後) by 降圧剤

T.2 = 5.5624, df1 = 3, df2 = 8, p-value = 0.02334

alternative hypothesis: true location difference is not equal to c(0,0,0)

```
> fit01 <- manova(cbind(投与前, 一時間後, 三時間後)~降圧剤, data=data01)
```

```
> summary(fit01, test="Wilks")
```

	Df	Wilks	approx F	num Df	den Df	Pr(>F)
降圧剤	1	0.32405	5.5624	3	8	0.02334 *
Residuals	10					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.3.2 Repeated ANOVA による分析

同じデータについて、repeated ANOVA による分析も行ってみることにしよう¹⁶。R のデフォルトの関数 `aov()` を用いることにすると¹⁷、従属変数が 1 列になるようデータの変形を行って、

```
> data02 <- data.frame(被験者=factor(rep(data01[,1],3)),
+                       降圧剤=rep(data01[,2],3),
+                       時間=c(rep("0h",12),rep("1h",12),rep("3h",12)),
+                       血圧=rbind(as.matrix(data01[,3]),
+                                   as.matrix(data01[,4])),
+                       )
```

¹⁶<https://ww2.coastal.edu/kingw/statistics/R-tutorials/repeated2.html> 参照。

¹⁷A が B にネストしている場合、`Error(B/A)` と指定する。

```

+                                     as.matrix(data01[,5]))
> head(data02,5)
  被験者 降圧剤 時間  血圧
1      1   10mg  0h   119
2      2   10mg  0h   110
3      3   10mg  0h   123
4      4   10mg  0h   130
5      5   10mg  0h   121
>
> fit02 <- aov(血圧~降圧剤*時間+Error(被験者/時間), data=data02)
> summary(fit02)

Error: 被験者
          Df Sum Sq Mean Sq F value Pr(>F)
降圧剤      1  182.2    182.2   1.806  0.209
Residuals 10 1009.4    100.9

Error: 被験者:時間
          Df Sum Sq Mean Sq F value  Pr(>F)
時間        2  742.1    371.0  10.343 0.000824 ***
降圧剤:時間  2  115.2     57.6   1.605 0.225663
Residuals   20  717.4     35.9

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

となる。時間の主効果は有意であるものの、降圧剤の主効果、降圧剤と時間の交互作用は有意でない。ただし、関数 `aov()` を用いるやり方では以下のような限界がある。

- タイプ I の平方和しか選べない（バランスデザインであることが大前提）
- 多重比較を `TukeyHSD()` で行うことができない
- 球面性の仮定の検証、自由度の調整ができない
- 単純主効果の検証（多重比較も含めて）ができない

パッケージ `car` の関数 `Anova()` を用いることで、平方和の種類と球面性の仮定については解決できる¹⁸。同じ分析を、タイプ III の平方和を指定して行った場合の結果は以下の通り。

```

> library(car)
> fit03 <- Anova(lm(as.matrix(data01[,3:5])~降圧剤, data=data01),
+               idata=data.frame(時間=factor(c(0,1,3))),
+               idesign=~時間, type="III")
> summary(fit03, multivariate=F)

```

¹⁸<https://egret.psychol.cam.ac.uk/statistics/R/anova.html>

Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

	SS	num Df	Error SS	den Df	F	Pr(>F)
(Intercept)	253235	1	1009.39	10	2508.7924	2.431e-13 ***
降圧剤	182	1	1009.39	10	1.8055	0.20874
時間	342	2	717.44	20	4.7685	0.02026 *
降圧剤:時間	115	2	717.44	20	1.6052	0.22566

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Mauchly Tests for Sphericity

	Test statistic	p-value
時間	0.56172	0.074618
降圧剤:時間	0.56172	0.074618

Greenhouse-Geisser and Huynh-Feldt Corrections for Departure from Sphericity

	GG eps	Pr(>F[GG])
時間	0.69528	0.03672 *
降圧剤:時間	0.69528	0.23315

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	HF eps	Pr(>F[HF])
時間	0.7721787	0.03157477
降圧剤:時間	0.7721787	0.23172547

この例では、Mauchly の検定により球面性の仮定は棄却されない。よって結果は変わらず、時間の主効果のみが有意ということになる。ただし、これでも TukeyHSD() で多重比較を行うことはできない。

6.3.3 対比を用いた線形性の検定

時間の主効果が有意であったので、時間に伴う血圧変化が線形であるか対比を用いて検定する。

$$C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \quad (6.20)$$

として、

$$\psi = XC' \quad (6.21)$$

対比を作る。

```

> contrast <- matrix(c(1,-1,0,
+                       0,1,-1), 2, 3, byrow=T)
> psi <- cbind(data01$投与前, data01$一時間後, data01$三時間後) %*% t(contrast)
> data03 <- cbind(data01[,1:2], data.frame(psi1=psi[,1], psi2=psi[,2]))
> head(data03, 5)
  被験者 降圧剤 psi1 psi2
1      1   10mg    6  -1
2      2   10mg   -5   3
3      3   10mg   -3  15
4      4   10mg    3  27
5      5   10mg    6  -9

```

これを ANOVA で分析することにする。

```

> data04 <- data.frame(被験者=factor(rep(data03[,1],2)), # ※ factor() を忘れないこと
+                      降圧剤=rep(data03[,2],2),
+                      対比=factor(c(rep(1,12),rep(2,12))),
+                      血圧変化=rbind(as.matrix(data03[,3]),
+                                      as.matrix(data03[,4])))
> head(data04,5)
  被験者 降圧剤 対比 血圧変化
1      1   10mg    1         6
2      2   10mg    1        -5
3      3   10mg    1        -3
4      4   10mg    1         3
5      5   10mg    1         6
> fit05 <- aov(血圧変化~降圧剤*対比+Error(被験者/対比), data=data04)
> summary(fit05)

```

Error: 被験者

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
降圧剤	1	2.7	2.67	0.055	0.819
Residuals	10	485.3	48.53		

Error: 被験者:対比

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
対比	1	48.2	48.2	0.692	0.4250
降圧剤:対比	1	337.5	337.5	4.847	0.0523 .
Residuals	10	696.3	69.6		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

対比の主効果は有意でなく、「血圧の減少は線形である」という帰無仮説は棄却されなかった ($F(1, 10) = 0.692, p = .425$)。なお、降圧剤の効果および交互作用も有意でない。

6.3.4 球面性の仮定とその検証

以下、「SAS による実験データの解析」第 12 章 (p.248-) による。

ある群での経時観測値 x_1, x_2, \dots, x_p の分散共分散行列が $\Sigma = \{\sigma_{ij}\}$ であるとする。このとき、正規性と分散共分散行列の等質性に加え、

$$\sigma_{ij} = \begin{cases} 2\alpha_i + \tau & (i = j) \\ \alpha_i + \alpha_j & (i \neq j) \end{cases} \quad \text{ただし } \tau > 0 \quad (6.22)$$

$$\iff V(x_i - x_j) = 2\tau \quad (\forall i \neq j) \quad (6.23)$$

でないといけない (Huynh & Feldt, 1970; **H-F 条件**)。

これは、 $1, 1, \dots, 1$ に直交する $(p-1)$ 本の正規直交対比 z_1, z_2, \dots, z_{p-1} について、その分散共分散行列が $\tau \mathbf{I}$ であることも同等である。これを、正規直交対比の分布に関する球面性の仮定という。

正規直交対比のプールされた平方和積和行列を \mathbf{A} とすると、

$$l = \frac{|\mathbf{A}|}{\text{tr}[\mathbf{A}/p]^p} \quad (6.24)$$

を **Mauchly** の基準といい ($0 \leq l \leq 1$)、 \mathbf{A} の自由度 df_A が大きければ、H-F 条件の下で

$$-\left(df_A - \frac{2p^2 + p + 2}{6p}\right) \log l \sim \chi_{\frac{1}{2}p(p+1)-1}^2 \quad (6.25)$$

となることが知られている。

6.3.5 球面性仮定からの逸脱と自由度修正

球面性の仮定が満たされていないと、 p 値は実際よりも小さくなることが知られている。

Greenhouse & Geiser (1959) により、母集団標準偏差 σ_{ij} の関数 ϵ

$$\frac{1}{p-1} \leq \epsilon \leq 1 \quad (6.26)$$

を用いて¹⁹、時間の主効果もしくは時間と被験者間要因との交互作用について

$$F = \frac{SS_1/(\epsilon \times df_1)}{SS_2/(\epsilon \times df_2)} \quad (6.27)$$

を自由度 $(\epsilon \times df_1, \epsilon \times df_2)$ の F 分布で近似できることが示された。

この ϵ について、定義式中の σ_{ij} を s_{ij} で置き換えたものを $\hat{\epsilon}_{GG}$ とすると、Huynh & Feldt (1976) によりバイアスを修正した

$$\hat{\epsilon}_{HF} = \frac{n(p-1)\hat{\epsilon}_{GG}-2}{(p-1)(n-1-(p-1)\hat{\epsilon}_{GG})} \quad (6.28)$$

が提案され、

$$\frac{1}{p-1} \leq \hat{\epsilon}_{GG} \leq \hat{\epsilon}_{HF} \quad (6.29)$$

であることが知られている。

球面性の仮定が棄却されたり、 $\hat{\epsilon}_{HF} < 1$ となるようなケースでは、自由度修正を行うことが推奨される。

¹⁹H-F 条件の下で $\epsilon = 1$