

## 一般化線形モデル (一元配置分散分析)

### 1 はじめに

分散分析は、質的変数（因子）が定めるカテゴリ（水準）の間で連続変数の平均値を比較する統計手法である。分散分析は、1 つまたはそれ以上の質的変数を説明変数とした正規線形モデルであり、デザイン行列  $\mathbf{X}$  はダミー変数から構成される。ダミー変数の定義の仕方はいくつかの方法があるが、もっとも適切な  $\mathbf{X}$  を選ぶことが重要である。

### 2 一元配置分散分析

作物重量の人工データである表 1 を考える。この実験では、対照、処理 A、処理 B の 3 つの成長条件について重量  $y_i$ （作物の乾燥重量）を比較している。反応変数となる乾燥重量は、3 つの水準から構成される 1 因子（成長条件）に依存している。つまり、各水準における乾燥重量の平均が異なるかどうかに関心がある。

表 1 3 つの異なる成長条件による作物の乾燥重量  $y_i$

	対照	処理 A	処理 B
	4.17	4.81	6.31
	5.58	4.17	5.12
	5.18	4.41	5.54
	6.11	3.59	5.50
	4.50	5.87	5.37
	4.61	3.83	5.29
	5.17	6.03	4.92
	4.53	4.89	6.15
	5.33	4.32	5.80
	5.14	4.69	5.26
$\sum y_i$	50.32	46.61	55.26
$\sum y_i^2$	256.27	222.92	307.13

一般に、 $J$  個の水準からなる因子に対して、各実験単位を  $J$  個の群のいずれかにランダムに割り当てるとき、**完全ランダム化実験**とよんでいる。データ形式は表 2 のようになる。

第  $j$  水準で観測された反応  $Y_{j1}, \dots, Y_{jn_j}$  はすべて同じ期待値をもち、**繰り返し**とよばれている。一般には、繰り返し数  $n_j$  は各水準に異なることが多いかもしれない。本講義では、議論を簡単にするために、すべての水準が同じ繰り返し数（サンプルサイズ）をもつ、つまり、 $n_j = K$  ( $j = 1, \dots, J$ ) と仮定し、反応  $\mathbf{y}$  を  $N = JK$  個の観測値からなる列ベクトルを次のように定義する。

$$\mathbf{y} = (Y_{11}, \dots, Y_{1K}, Y_{21}, \dots, Y_{2K}, \dots, Y_{J1}, \dots, Y_{JK})^\top$$

因子水準間で反応の平均値が異なるという仮説を検定するためのモデルを異なる視点から 3 つ考

える。

表 2  $J$  個の水準からなる因子 A に関する完全ランダム化実験のデータ

因子水準				
	$A_1$	$A_2$	$\cdots$	$A_J$
$Y_{11}$	$Y_{11}$	$Y_{21}$	$\cdots$	$Y_{J1}$
$Y_{12}$	$Y_{12}$	$Y_{22}$	$\cdots$	$Y_{J2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Y_{1n_1}$	$Y_{1n_1}$	$Y_{2n_2}$	$\cdots$	$Y_{Jn_J}$
計	$Y_{1\cdot}$	$Y_{2\cdot}$	$\cdots$	$Y_{J\cdot}$

## 2.1 モデル 1

もっとも単純な単純なモデルは、次のように表すことができる。

$$E[Y_{jk}] = \mu_j \quad (j = 1, \dots, J) \quad (1)$$

正規線形モデルであることを意識して、より統一的な表現をすれば、次のようになる。

$$E[Y_i] = \sum_{j=1}^J x_{ij} \mu_j \quad (i = 1, \dots, N)$$

ただし、 $Y_i$  が第  $j$  水準に属する反応であれば  $x_{ij} = 1$  とし、それ以外の  $j$  については  $x_{ij} = 0$  とする。これを行列形式で表すと次のようになる。

$$E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$$

ただし、

$$\boldsymbol{\beta} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_J \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix}$$

ここで、 $\mathbf{0}$  は要素がすべて 0 からなる  $K \times 1$  ベクトル、 $\mathbf{1}$  は要素がすべて 1 からなる  $K \times 1$  ベクトルである。 $\mathbf{X}$  は  $N \times J$  行列であり、 $\mathbf{X}^\top \mathbf{X}$  は  $J \times J$  行列である。

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} K & 0 & \cdots & 0 \\ 0 & K & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & K \end{pmatrix} = K \mathbf{I}_J, \quad \mathbf{X}^\top \mathbf{y} = \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix}$$

$\mathbf{X}^\top \mathbf{X}$  は正則であり、その逆行列  $(\mathbf{X}^\top \mathbf{X})^{-1}$  は次のようになる。

$$(\mathbf{X}^\top \mathbf{X})^{-1} = \frac{1}{K} \mathbf{I}_J$$

したがって、 $\beta$  の最尤推定量は次のようになる。

$$\hat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} = \frac{1}{K} \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix} = \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_J \end{pmatrix}$$

回帰平方和  $\hat{\beta}^\top \mathbf{X}^\top \mathbf{y}$  は次のようになる。

$$\hat{\beta}^\top \mathbf{X}^\top \mathbf{y} = \frac{1}{K} (Y_{1\cdot}, Y_{2\cdot}, \dots, Y_{J\cdot}) \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix} = \frac{1}{K} \sum_{j=1}^J Y_{j\cdot}^2$$

当てはめ値は次のようになる。

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{X} \hat{\beta} \\ &= \begin{pmatrix} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix} \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_J \end{pmatrix} \\ &= (\bar{Y}_1, \dots, \bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_2, \dots, \bar{Y}_J, \dots, \bar{Y}_J)^\top \end{aligned}$$

(1) 式のモデルでは、2 つ以上の因子が反応に関係する場合に、各因子の水準（または、それらの組み合わせ）における反応値が、平均的あるいは特定の反応値からどれだけ乖離しているか、といった観点を検討できないという欠点がある。そこで、そのようなことがパラメータの値に反映されるモデル化を考える。

## 2.2 モデル 2

各因子の水準における反応値が平均的な反応値からどれだけ乖離しているか検討するために、次のモデルを考える。

$$E[Y_{jk}] = \mu + \alpha_j \quad (j = 1, \dots, J) \quad (2)$$

ここで、 $\mu$  は全水準に対する共通の平均効果、 $\alpha_j$  は水準  $j$  における固有の付加的な効果を表す。

(2) 式のモデルでは、パラメータの総数は  $J + 1$  個であり、行列形式で表すと次のようになる。

$$\beta = \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_J \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & & \mathbf{0} \\ \vdots & \vdots & & \ddots & \vdots \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix}$$

ここで、 $\mathbf{0}$  は要素がすべて 0 からなる  $K \times 1$  ベクトル、 $\mathbf{1}$  は要素がすべて 1 からなる  $K \times 1$  ベクトルである。 $\mathbf{X}$  は  $N \times (J + 1)$  行列であり、 $\mathbf{X}^\top \mathbf{X}$  は  $(J + 1) \times (J + 1)$  行列である。

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} N & K & K & \cdots & K \\ K & K & 0 & \cdots & 0 \\ K & 0 & K & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ K & 0 & \cdots & 0 & K \end{pmatrix}, \quad \mathbf{X}^\top \mathbf{y} = \begin{pmatrix} Y_{\cdot\cdot} \\ Y_{1\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix}$$

$\mathbf{X}^\top \mathbf{X}$  の第 1 行目 (第 1 列目) は, 他の行 (列) の和になっているため, 特異 (非正則) 行列である。したがって, 正規方程式  $\mathbf{X}^\top \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^\top \mathbf{y}$  の解は一意に定まらない。一般解は,  $\lambda$  を任意の定数として, 次のように表せる。

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_J \end{pmatrix} = \frac{1}{K} \begin{pmatrix} 0 \\ Y_{1\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix} - \lambda \begin{pmatrix} -1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

ここで, **零和制約** (sum-to-zero constraint)

$$\sum_{j=1}^J \alpha_j = 0$$

を仮定すれば,

$$\frac{1}{K} \sum_{j=1}^K Y_{j\cdot} - J\lambda = 0$$

となることから, 次式を得る。

$$\lambda = \frac{1}{JK} \sum_{j=1}^K Y_{j\cdot} = \frac{Y_{\cdot\cdot}}{N}$$

したがって, 正規方程式の解は次のようになる。

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_J \end{pmatrix} = \begin{pmatrix} \frac{Y_{\cdot\cdot}}{N} \\ \frac{1}{K} Y_{1\cdot} - \frac{Y_{\cdot\cdot}}{N} \\ \vdots \\ \frac{1}{K} Y_{J\cdot} - \frac{Y_{\cdot\cdot}}{N} \end{pmatrix}$$

回帰平方和  $\hat{\boldsymbol{\beta}}^\top \mathbf{X}^\top \mathbf{y}$  は次のようになる。

$$\begin{aligned} \hat{\boldsymbol{\beta}}^\top \mathbf{X}^\top \mathbf{y} &= \left( \frac{Y_{\cdot\cdot}}{N}, \frac{1}{K} Y_{1\cdot} - \frac{Y_{\cdot\cdot}}{N}, \dots, \frac{1}{K} Y_{J\cdot} - \frac{Y_{\cdot\cdot}}{N} \right) \begin{pmatrix} Y_{\cdot\cdot} \\ Y_{1\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix} \\ &= \frac{Y_{\cdot\cdot}^2}{N} + \sum_{j=1}^J \left( \frac{1}{K} Y_{j\cdot} - \frac{Y_{\cdot\cdot}}{N} \right) Y_{j\cdot} \\ &= \frac{Y_{\cdot\cdot}^2}{N} + \frac{1}{K} \sum_{j=1}^J Y_{j\cdot}^2 - \frac{Y_{\cdot\cdot}}{N} \sum_{j=1}^J Y_{j\cdot} \\ &= \frac{1}{K} \sum_{j=1}^J Y_{j\cdot}^2 \end{aligned}$$

これは, (1) 式のモデルの回帰平方和と一致することがわかる。

当てはめ値は次のようになる。

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & & \mathbf{0} \\ \vdots & \vdots & & \ddots & \vdots \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix} \begin{pmatrix} \frac{Y_{..}}{N} \\ \frac{1}{K}Y_{1.} - \frac{Y_{..}}{N} \\ \vdots \\ \frac{1}{K}Y_{J.} - \frac{Y_{..}}{N} \end{pmatrix} \\ &= (\bar{Y}_1, \dots, \bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_2, \dots, \bar{Y}_J, \dots, \bar{Y}_J)^\top\end{aligned}$$

これは、(1) 式のモデルの当てはめ値と一致することがわかる。

## 2.3 モデル 3

各因子の水準における反応値が特定の反応値からどれだけ乖離しているか検討するために、次のモデルを考える。

$$E[Y_{jk}] = \mu + \alpha_j \quad (j = 1, \dots, J) \quad (3)$$

ただし、 $\alpha_1 = 0$  とする。このモデルのもとで、 $\mu$  は第 1 水準での平均効果、 $\alpha_j$  ( $j = 2, \dots, J$ ) は第 1 水準と第  $j$  水準との差を表すパラメータと考えることができる。これは、**端点制約によるパラメータ化** (corner-point parameterization) とよばれている。

このモデルを行列形式で表すと次のようになる。

$$E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$$

ただし、

$$\boldsymbol{\beta} = \begin{pmatrix} \mu \\ \alpha_2 \\ \vdots \\ \alpha_J \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{1} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix}$$

ここで、 $\mathbf{0}$  は要素がすべて 0 からなる  $K \times 1$  ベクトル、 $\mathbf{1}$  は要素がすべて 1 からなる  $K \times 1$  ベクトルである。 $\mathbf{X}$  は  $N \times J$  行列であり、 $\mathbf{X}^\top \mathbf{X}$  は  $J \times J$  行列である。

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} N & K & K & \cdots & K \\ K & K & 0 & \cdots & 0 \\ K & 0 & K & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ K & 0 & \cdots & 0 & K \end{pmatrix}, \quad \mathbf{X}^\top \mathbf{y} = \begin{pmatrix} Y_{..} \\ Y_{2.} \\ \vdots \\ Y_{J.} \end{pmatrix}$$

ただし、 $Y_{..} = \sum_{j=1}^J Y_{j.}$  である。 $\mathbf{X}^\top \mathbf{X}$  は正則であり、その逆行列  $(\mathbf{X}^\top \mathbf{X})^{-1}$  は次のようになる。

$$(\mathbf{X}^\top \mathbf{X})^{-1} = \frac{1}{K} \begin{pmatrix} 1 & -1 & -1 & \cdots & -1 \\ -1 & 2 & 1 & \cdots & 1 \\ -1 & 1 & 2 & & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ -1 & 1 & \cdots & 1 & 2 \end{pmatrix}$$

したがって、 $\beta$  の最尤推定量は次のようになる。

$$\hat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} = \frac{1}{K} \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} - Y_{1\cdot} \\ \vdots \\ Y_{J\cdot} - Y_{1\cdot} \end{pmatrix}$$

回帰平方和  $\hat{\beta}^\top \mathbf{X}^\top \mathbf{y}$  は次のようになる。

$$\begin{aligned} \hat{\beta}^\top \mathbf{X}^\top \mathbf{y} &= \frac{1}{K} (Y_{1\cdot}, Y_{2\cdot} - Y_{1\cdot}, \dots, Y_{J\cdot} - Y_{1\cdot}) \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} \\ \vdots \\ Y_{J\cdot} \end{pmatrix} \\ &= \frac{1}{K} \left[ Y_{1\cdot} Y_{1\cdot} + \sum_{j=2}^J (Y_{j\cdot} - Y_{1\cdot}) Y_{j\cdot} \right] \\ &= \frac{1}{K} \left[ \sum_{j=1}^J Y_{1\cdot} Y_{j\cdot} + \sum_{j=2}^J Y_{j\cdot}^2 - \sum_{j=2}^J Y_{1\cdot} Y_{j\cdot} \right] \\ &= \frac{1}{K} \sum_{j=1}^J Y_{j\cdot}^2 \end{aligned}$$

これは、(1) 式のモデルの回帰平方和と一致することがわかる。

当てはめ値は次のようになる。

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{X} \hat{\beta} \\ &= \frac{1}{K} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 1 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} - Y_{1\cdot} \\ \vdots \\ Y_{J\cdot} - Y_{1\cdot} \end{pmatrix} \\ &= (\bar{Y}_1, \dots, \bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_2, \dots, \bar{Y}_J, \dots, \bar{Y}_J)^\top \end{aligned}$$

これは、(1) 式のモデルの当てはめ値と一致することがわかる。

## 2.4 逸脱度

(1) から (3) のモデルはパラメータ化の方法こそ異なるが、いずれも同じ平方和  $\hat{\beta}^\top \mathbf{X}^\top \mathbf{y}$  を与えることがわかる。したがって、(1) から (3) のモデルの逸脱度は次のようになる。

$$D_1 = \frac{1}{\sigma^2} (\mathbf{y}^\top \mathbf{y} - \hat{\beta}^\top \mathbf{X}^\top \mathbf{y}) = \frac{1}{\sigma^2} \left( \sum_{j=1}^J \sum_{k=1}^K Y_{jk}^2 - \frac{1}{K} \sum_{j=1}^J Y_{j\cdot}^2 \right)$$

(1) から (3) のモデルは、すべて「各水準間で反応平均が異なる」という仮説  $H_1$  に対応している。「反応平均はすべて等しい」という帰無仮説  $H_0$  と比較するために、 $H_0$  のもとでのモデル  $E[Y_{jk}] = \mu$  を考える。このとき、 $\beta = \mu$  であり、 $\mathbf{X}$  はすべての要素が 1 からなる  $N \times 1$  ベクトルである。つまり、 $\mathbf{X}^\top \mathbf{X} = N$  かつ  $\mathbf{X}^\top \mathbf{y} = Y_{1\cdot}$  となることから、パラメータの推定量は次のよう

になる。

$$\hat{\beta} = \hat{\mu} = \frac{Y_{..}}{N}$$

回帰平方和は次のようになる。

$$\hat{\beta} \mathbf{X}^\top \mathbf{y} = \frac{Y_{..}^2}{N}$$

したがって、 $H_0$  のもとでのモデル  $E[Y_{jk}] = \mu$  の逸脱度は次のようになる。

$$D_0 = \frac{1}{\sigma^2} \left( \mathbf{y}^\top \mathbf{y} - \hat{\beta}^\top \mathbf{X}^\top \mathbf{y} \right) = \frac{1}{\sigma^2} \left( \sum_{j=1}^J \sum_{k=1}^K Y_{jk}^2 - \frac{Y_{..}^2}{N} \right)$$

## 2.5 仮説検定

帰無仮説  $H_0$  の検定を考える。ここで、 $H_1$  は正しく、 $D_1 \sim \chi^2(N - J)$  と仮定してよいものとする。もし  $H_0$  が正しければ、さらに  $D_0 \sim \chi^2(N - 1)$  と仮定できる。 $H_0$  が正しくなければ、 $D_0$  は非心カイ二乗分布に従う。したがって、 $H_0$  のもとでは、

$$D_0 - D_1 = \frac{1}{\sigma^2} \left( \frac{1}{K} \sum_{j=1}^J Y_{j.}^2 - \frac{Y_{..}^2}{N} \right) \sim \chi^2(J - 1)$$

となることから、

$$F = \frac{\frac{D_0 - D_1}{J - 1}}{\frac{D_1}{N - J}} \sim F(J - 1, N - J)$$

となる。もし  $H_0$  が正しくなければ、 $F$  値は自由度  $(J - 1, N - J)$  の  $F$  分布に比べて大きくなる。以上の仮説検定の議論は、通常、分散分析表にまとめられる。

## 参考文献

- [1] 江金芳. (2016). 一般化線形モデル. 朝倉書店.
- [2] Dobson, A. J and Barnett, A. G. (2018). *An introduction to generalized linear models, 4th Edition*. Chapman and Hall/CRC.