STAT 500

Estimability

Identifiability and Estimability

Definition: For a linear model $E(\mathbf{Y}) = X\boldsymbol{\beta}$, the parameter vector $\boldsymbol{\beta}$ is *identifiable* if $X\boldsymbol{\beta}_1 = X\boldsymbol{\beta}_2$ implies $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$

- Only *identifiable* parameters can be estimated
- Linear functions of identifiable parameters are called *estimable*
- Unbiased estimators can be found for estimable functions of model parameters

Example: Identifiability

One-Way ANOVA Cell Means Model

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{21} \\ Y_{22} \\ Y_{23} \\ Y_{24} \\ Y_{31} \\ Y_{32} \\ Y_{33} \\ Y_{34} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \\ \epsilon_{24} \\ \epsilon_{31} \\ \epsilon_{32} \\ \epsilon_{33} \\ \epsilon_{34} \end{bmatrix}$$

Example: Identifiability

Let $\boldsymbol{\beta}_1 = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$

$$\begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{13}) \\ E(Y_{14}) \\ E(Y_{21}) \\ E(Y_{22}) \\ E(Y_{23}) \\ E(Y_{24}) \\ E(Y_{31}) \\ E(Y_{32}) \\ E(Y_{33}) \\ E(Y_{34}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_1 \\ \beta_1 \\ \beta_1 \\ \beta_2 \\ \beta_2 \\ \beta_2 \\ \beta_2 \\ \beta_3 \\ \beta_3 \\ \beta_3 \\ \beta_3 \end{bmatrix}$$

Example: Identifiability

Let $\boldsymbol{\beta}_2 = \begin{pmatrix} \beta_1^* \\ \beta_2^* \\ \beta_3^* \end{pmatrix}$

$$\begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{13}) \\ E(Y_{14}) \\ E(Y_{21}) \\ E(Y_{22}) \\ E(Y_{23}) \\ E(Y_{24}) \\ E(Y_{31}) \\ E(Y_{32}) \\ E(Y_{33}) \\ E(Y_{34}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1^* \\ \beta_2^* \\ \beta_3^* \end{bmatrix} = \begin{bmatrix} \beta_1^* \\ \beta_1^* \\ \beta_1^* \\ \beta_1^* \\ \beta_2^* \\ \beta_2^* \\ \beta_2^* \\ \beta_2^* \\ \beta_3^* \\ \beta_3^* \\ \beta_3^* \\ \beta_3^* \end{bmatrix}$$

Example: Identifiability

One-Way ANOVA Cell Means Model

- For $X\boldsymbol{\beta}_1 = X\boldsymbol{\beta}_2$, we must have $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$
- For this model, $\boldsymbol{\beta}$ is identifiable.
- The vector of response means uniquely determines the values of the parameter vector $\boldsymbol{\beta}$.

Example: Identifiability

One-Way ANOVA Effects Model

$$
\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{21} \\ Y_{22} \\ Y_{23} \\ Y_{24} \\ Y_{31} \\ Y_{32} \\ Y_{33} \\ Y_{34} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \\ \epsilon_{24} \\ \epsilon_{31} \\ \epsilon_{32} \\ \epsilon_{33} \\ \epsilon_{34} \end{bmatrix}
$$

Example: Identifiability

Let $\boldsymbol{\beta}_1 = \begin{pmatrix} \mu_3 \\ \mu_1 - \mu_3 \\ \mu_2 - \mu_3 \\ 0 \end{pmatrix}$

$$\begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{13}) \\ E(Y_{14}) \\ E(Y_{21}) \\ E(Y_{22}) \\ E(Y_{23}) \\ E(Y_{24}) \\ E(Y_{31}) \\ E(Y_{32}) \\ E(Y_{33}) \\ E(Y_{34}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_3 \\ \mu_1 - \mu_3 \\ \mu_2 - \mu_3 \\ 0 \end{bmatrix}$$

Example: Identifiability

Let $\boldsymbol{\beta}_2 = \begin{pmatrix} 0 \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$

$$\begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{13}) \\ E(Y_{14}) \\ E(Y_{21}) \\ E(Y_{22}) \\ E(Y_{23}) \\ E(Y_{24}) \\ E(Y_{31}) \\ E(Y_{32}) \\ E(Y_{33}) \\ E(Y_{34}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

Example: Identifiability

One-Way ANOVA Effects Model

- $X\boldsymbol{\beta}_1 = X\boldsymbol{\beta}_2$ but $\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$, so $\boldsymbol{\beta}$ is not identifiable.
- The vector of response means does not uniquely determine the values of the parameter vector $\boldsymbol{\beta}$.

Estimable Functions

An estimable function is a linear function of identifiable parameters

- Estimable functions are reasonable things to estimate
- Estimable functions have the same interpretation regardless of the constraints placed on the parameters to get a solution to the normal equations
- Least squares estimates of estimable functions are not affected by the choice of constraints placed on the parameters to get a solution to the normal equations

Estimable Functions

Definition: For a linear model $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, we say that a linear function of $\boldsymbol{\beta}$, $C\boldsymbol{\beta}$, is *estimable* if

$$C\boldsymbol{\beta} = AX\boldsymbol{\beta} = AE(\mathbf{Y})$$

for some matrix A.

Note that:

- C is a $m \times k$ matrix of constants that defines m estimable linear functions of the parameters, and $C = AX$ for some A

Estimable Functions: Example

One Way ANOVA Effects model

$$\begin{bmatrix} E(Y_{11}) \\ E(Y_{12}) \\ E(Y_{13}) \\ E(Y_{14}) \\ E(Y_{21}) \\ E(Y_{22}) \\ E(Y_{23}) \\ E(Y_{24}) \\ E(Y_{31}) \\ E(Y_{32}) \\ E(Y_{33}) \\ E(Y_{34}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_3 \\ \mu + \alpha_3 \\ \mu + \alpha_3 \\ \mu + \alpha_3 \end{bmatrix}$$

Estimable Functions: Example

Show that $\mu + \alpha_1 = [1\ 1\ 0\ 0]\boldsymbol{\beta}$ is estimable.

Let $A = \left[\frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\right]$

$$\begin{aligned} AX\boldsymbol{\beta} = AE(\mathbf{Y}) &= \frac{1}{4}E(Y_{11}) + \frac{1}{4}E(Y_{12}) + \frac{1}{4}E(Y_{13}) + \frac{1}{4}E(Y_{14}) \\ &= \frac{1}{4}(\mu + \alpha_1) + \frac{1}{4}(\mu + \alpha_1) + \frac{1}{4}(\mu + \alpha_1) + \frac{1}{4}(\mu + \alpha_1) \\ &= \mu + \alpha_1 \end{aligned}$$

Alternatively, let $A = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$

$$AX\boldsymbol{\beta} = AE(\mathbf{Y}) = E(Y_{11}) = \mu + \alpha_1$$

Estimable Functions: Example

Show that $\begin{bmatrix} \mu + \alpha_2 \\ \mu + \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{\beta}$ is estimable.

Let $A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

$$AX\boldsymbol{\beta} = AE(\mathbf{Y}) = \begin{bmatrix} E(Y_{21}) \\ E(Y_{31}) \end{bmatrix} = \begin{bmatrix} \mu + \alpha_2 \\ \mu + \alpha_3 \end{bmatrix}$$

Estimable Functions: Example

Show that $\alpha_1 - \alpha_2 = [0\ 1\ \ -1\ 0]\boldsymbol{\beta}$ is estimable

$$
\begin{aligned}
\alpha_1 - \alpha_2 &= (\mu + \alpha_1) - (\mu + \alpha_2) \\
&= E(Y_{11}) - E(Y_{21}) \\
&= [1\ 0\ 0\ 0\ \text{-1}\ 0\ 0\ 0\ 0\ 0\ 0\ 0]E(\mathbf{Y}) \\
&= AE(\mathbf{Y})
\end{aligned}
$$

Estimable Functions: Example

Show that $\begin{bmatrix} \alpha_2 - \alpha_3 \\ 2\mu + 3\alpha_1 - \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & -1 \\ 2 & 3 & -1 & 0 \end{bmatrix} \boldsymbol{\beta}$ is estimable

$$\begin{aligned} \begin{bmatrix} \alpha_2 - \alpha_3 \\ 2\mu + 3\alpha_1 - \alpha_2 \end{bmatrix} &= \begin{bmatrix} (\mu + \alpha_2) - (\mu + \alpha_3) \\ 3(\mu + \alpha_1) - (\mu + \alpha_2) \end{bmatrix} \\ &= \begin{bmatrix} E(Y_{21}) - E(Y_{31}) \\ 3E(Y_{11}) - E(Y_{21}) \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_3 \\ \mu + \alpha_3 \\ \mu + \alpha_3 \\ \mu + \alpha_3 \end{bmatrix}$$

$$= AE(\mathbf{Y})$$

Functions that are Not Estimable

$$\mu,\ \alpha_1,\ \alpha_2,\ \alpha_3,\ 3\alpha_1,\ \alpha_1+\alpha_2$$

To show that

$$c^T\boldsymbol{\beta} = c_1\mu + c_2\alpha_1 + c_3\alpha_2 + c_4\alpha_3$$

is <u>not</u> estimable, one must show that there is no non-random matrix A for which

$$c^T\boldsymbol{\beta} = c_1\mu + c_2\alpha_1 + c_3\alpha_2 + c_4\alpha_3 = AX\boldsymbol{\beta} = AE(\mathbf{Y})$$

Non-Estimable Functions: Example

Show $\alpha_1 = [0\ 1\ 0\ 0]\boldsymbol{\beta}$ is not estimable.

For α_1 to be estimable, we would need to find a matrix A such that

$$\begin{aligned}
\alpha_1 &= AE(\mathbf{Y}) \\
&= a_1E(Y_{11}) + a_2E(Y_{12}) + a_3E(Y_{13}) + a_4E(Y_{14}) \\
&\quad + a_5E(Y_{21}) + a_6E(Y_{22}) + a_7E(Y_{23}) + a_8E(Y_{24}) \\
&\quad + a_9E(Y_{31}) + a_{10}E(Y_{32}) + a_{11}E(Y_{33}) + a_{12}E(Y_{34}) \\
&= (a_1 + a_2 + a_3 + a_4)(\mu + \alpha_1) \\
&\quad + (a_5 + a_6 + a_7 + a_8)(\mu + \alpha_2) \\
&\quad + (a_9 + a_{10} + a_{11} + a_{12})(\mu + \alpha_3)
\end{aligned}$$

Non-Estimable Functions: Example

This implies that

$$0 = (a_5 + a_6 + a_7 + a_8) = (a_9 + a_{10} + a_{11} + a_{12})$$

and

$$\begin{aligned} \alpha_1 &= (a_1 + a_2 + a_3 + a_4)(\mu + \alpha_1) \\ &= (a_1 + a_2 + a_3 + a_4)\mu + (a_1 + a_2 + a_3 + a_4)\alpha_1 \end{aligned}$$

This is not possible, so α_1 is not estimable.

Rules for Estimable Functions

For a linear model $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$

- The expectation of any observation is estimable.
- A linear combination of estimable functions is estimable.
- Each element of $\boldsymbol{\beta}$ is estimable if and only if rank$(X) = k =$ number of columns in X.
- Every $\mathbf{c}^T\boldsymbol{\beta}$ is estimable if and only if rank$(X) = k =$ number of columns in X.
- Let X_j be the j-th column of X. β_j is not estimable if and only if $X_j = \sum_{j \neq l} c_l X_l$ for some set of scalars $\{\ c_i :\ \ j \neq l\}$.

Estimable Functions: Example

- Multiple Linear Regression - show β_j is estimable for all j

  For a multiple linear regression model, the design matrix X is typically full rank (no perfect correlation among predictors). So every element of $\boldsymbol{\beta}$ is estimable.

- Multiple Linear Regression - show $\mu_{Y|x_1,x_2,\ldots,x_k}$ is estimable

  Since every element of $\boldsymbol{\beta}$ is estimable, a linear combination of $\boldsymbol{\beta}$ is also estimable.

  $$\mu_{Y|x_1,x_2,\ldots,x_k} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

Non-Estimable Functions: Example

- One-Way ANOVA Effects Model - show μ is not estimable.

  Let X_j be the jth column in the design matrix X.

  $\mu = \beta_1$ is not estimable since we can write

$$X_1 = X_2 + X_3 + X_4 = 1 * X_2 + 1 * X_3 + 1 * X_4$$

Estimable Functions

For a linear model

$$\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- Definitions of estimable functions of the elements of the parameter vector $\boldsymbol{\beta}$ depend on the linear model for the expected responses $E(\mathbf{Y}) = X\boldsymbol{\beta}$.

- No assumption is made about $Var(\mathbf{Y})$ or $Var(\boldsymbol{\epsilon})$ or the shape of the distribution of $\mathbf{Y}$ or $\boldsymbol{\epsilon}$.

Least Squares Estimator for Estimable Functions

Let $\mathbf{b} = (X^T X)^- X^T \mathbf{Y}$ be a solution to the normal equations. For any estimable function $C\boldsymbol{\beta}$, the least squares estimator for this estimable function is $C\mathbf{b}$ and is unique.

This means that $C\mathbf{b}$ has the same value regardless of

- constraints placed on parameters
- the choice of the generalized inverse matrix

The Gauss-Markov Theorem

Theorem For the Gauss-Markov model, $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, with

$$E(\mathbf{Y}) = X\boldsymbol{\beta} \text{ and } Var(\mathbf{Y}) = \sigma^2 I,$$

the OLS estimator $C\mathbf{b}$ of an estimable function $C\boldsymbol{\beta}$ is the unique best linear unbiased estimator (BLUE) of this estimable function.

'Best' means out of all possible linear unbiased estimators, the one with the smallest variance.