HW6 MATH 7570 Fall 2023

Wenjuan Bian

October 2023

1. In the textbook, Section 6.3, problem 3.
2. In the textbook, Section 6.3, problem 9.
3. In the textbook, Section 6.3, problem 13.
4. In the textbook, Section 6.3, problem 17.
5. In the textbook, Section 6.3, problem 19.
6. In the textbook, Section 6.3, problem 23.

1. T3 Prove Corollary 6.1.2.2. Corollary 6.1.2.2 The elements of $X\beta$ are estimable under the model $\{y, X\beta\}$; in fact, those elements span the space of estimable functions for that model

   Solution:

   The i^{th} element of $X\beta$ is $x_i^T\beta$, where x_i^T is the i^{th} row of X. Therefore, $x_i^T \in \mathcal{R}(X)$, so the i^{th} element of $X\beta$ is estimable. Because $\text{span}(x_1^T, \ldots, x_n^T) = \mathcal{R}(X)$, the elements of $X\beta$ span the space of estimable functions.

2. T9 Determine $\mathcal{N}(X)$ for the one-factor model and for the two-way main effects model with no empty cells, and then use Corollary 6.1.2.5 to obtain the same characterizations of estimable functions given in Examples 6.1-2 and 6.1-3 for these two models.

   Solution:

   For the One-Factor Model:

   The distinctive rows of the matrix X is $(1_q, I_q)$. Thus

   $$N(X) = \{v : v = a(-1)1_q,\ a \in \mathbb{R}\}.$$

   According to Corollary 6.1.2.5, the expression $c^T\beta$ is estimable precisely when c is orthogonal to $\mathcal{N}(X)$. This implies:

   $$c^T a(-1)1_q = 0\ \forall\ a \in \mathbb{R},$$

That is, $c_1 = -\sum_{i=2}^{q} c_i$.

For the Two-Way Main Effects Model (without Empty Cells):

The distinctive rows of the matrix X for this model are given by $(1_{qm}, I_q \otimes 1_m, 1_q \otimes I_m)$. From this, it follows that:

$$N(X) = \{v : v = (a, b1_q^T, -(a+b)1_m^T)^T,\ a, b \in \mathbb{R}\}.$$

By referring again to Corollary 6.1.2.5, the term $c^T\beta$ is estimable if:

$$c^T \begin{pmatrix} a \\ b1_q \\ -(a+b)1_m \end{pmatrix} = 0 \ \forall a, b \in \mathbb{R}.$$

Upon expansion, this condition transforms to:

$$ac_1 + b\left(\sum_{i=2}^{q+1} c_i\right) - (a+b)\left(\sum_{i=q+2}^{q+m+1} c_i\right) = 0$$

from which we can infer:

$$c_1 = \sum_{i=2}^{q+1} c_i = \sum_{i=q+2}^{q+m+1} c_i$$

.

3. T13 Consider the following two-way layout, where the number in each cell indicates how many observations are in that cell:

	Levels of B									
Levels of A	1	2	3	4	5	6	7	8	9	10
1		1						2		
2	1									
3			4	1			1			
4		2			2					
5					1	1				
6				1						
7						1		2		
8							1		3	
9										2

Consider the main effects model:

$$y_{ijk} = \mu + \alpha_i + \gamma_j + e_{ijk} \quad (i = 1, \ldots, 9; j = 1, \ldots, 10)$$

for these observations (where k indexes the observations, if any, within cell (i, j) of the table).

(a) Which α-contrasts and γ -contrasts are estimable?

(b) Give a basis for the set of all estimable functions.

(c) What is the rank of the associated model matrix X?

Solution:

1. The "3+e" procedure, followed by an appropriate rearrangement of rows and columns, yields the subsequent set of disconnected subarrays:

	Levels of B									
Levels of A	4	5	6	10	3	7	8	2	9	1
3	4	e	1	1						
4	e	2	e	2						
6	e	1	e	e						
7	1	2	e	e						
1					1	2	e			
5					e	1	1			
2								e	1	
8								1	3	
9										2

From the table, the estimable α-contrasts are:

$$\sum_{i=1}^{9} d_i \alpha_i$$

where

$$\sum_{i \in \{3,4,6,7\}} d_i = \sum_{i \in \{1,5\}} d_i = \sum_{i \in \{2,8\}} d_i = d_9 = 0$$

and the estimable γ-contrasts are of the form:

$$\sum_{j=1}^{10} g_j \gamma_j$$

where

$$\sum_{j\in\{4,5,6,10\}} g_j = \sum_{j\in\{3,7,8\}} g_j = \sum_{j\in\{2,9\}} g_j = g_1 = 0$$

.

2.

$$\{\mu+\alpha_3+\gamma_4, \alpha_4-\alpha_3, \alpha_6-\alpha_3, \alpha_7-\alpha_3, \gamma_5-\gamma_4, \gamma_6-\gamma_4, \gamma_{10}-\gamma_4, \mu+\alpha_1+\gamma_3, \alpha_5-\alpha_1, \gamma_7-\gamma_3, \gamma_8-\gamma_3, \mu+\alpha_2+$$

3.

$$\text{rank}(X) = 15$$

as deduced from part (b).

4. T17 Consider the three-way main effects model

$$y_{ijkl} = \mu + \alpha_i + \gamma_j + \delta_k + e_{ijkl}$$

where $(i = 1, \dots, q; j = 1, \dots, m; k = 1, \dots, s; l = 1, \dots, n_{ijk})$, and suppose that there are no empty cells (i.e., $n_{ijk} > 0$ for all i, j, k).

(a) What conditions must the elements of c satisfy for $c^T\beta$ to be estimable?

Solution:

By Theorem 6.1.2, $c^T\beta$ is estimable if and only if the coefficients on the α_i's sum to the coefficient on μ and the coefficients on the γ_j's and δ_k's sum to the coefficient on μ as well.

(b) Find a basis for the set of all estimable functions.

Solution:

A basis for the set of all estimable functions is:

$$\{\mu+\alpha_1+\gamma_1+\delta_1, \alpha_2-\alpha_1, \alpha_3-\alpha_1, \dots, \alpha_q-\alpha_1, \gamma_2-\gamma_1, \dots, \gamma_m-\gamma_1, \delta_2-\delta_1, \dots, \delta_s-\delta_1\}$$

(c) What is the rank of the model matrix X?

Solution

By part (b) and Corollary 6.1.2.3,

$$\text{rank}(X) = 1 + (q-1) + (m-1) + (s-1) = q + m + s - 2$$

5. T19. Consider a three-factor crossed classification in which Factors A, B, and C each have two levels, and suppose that there is exactly one observation in the following four (of the eight) cells: 111, 211, 122, 222. The other cells are empty. Consider the following main effects model for these observations:

$$y_{ijk} = \mu + \alpha_i + \gamma_j + \delta_k + e_{ijk}$$

where $(i, j, k) \in \{(1, 1, 1), (2, 1, 1), (1, 2, 2), (2, 2, 2)\}$.

   (a) Which of the functions $\alpha_2 - \alpha_1$, $\gamma_2 - \gamma_1$, and $\delta_2 - \delta_1$ are estimable?
   Solution:
   By Corollary 6.1.2.1, $\alpha_2 - \alpha_1$ is estimable because

$$-\frac{1}{2}(\mu + \alpha_1 + \gamma_1 + \delta_1) + \frac{1}{2}(\mu + \alpha_2 + \gamma_1 + \delta_1) - \frac{1}{2}(\mu + \alpha_1 + \gamma_2 + \delta_2) + \frac{1}{2}(\mu + \alpha_2 + \gamma_2 + \delta_2)$$
$$= \alpha_2 - \alpha_1.$$

   No linear combination of the same four functions yields either $\gamma_2 - \gamma_1$ or $\delta_2 - \delta_1$, so neither $\gamma_2 - \gamma_1$ nor $\delta_2 - \delta_1$ is estimable.

   (b) Find a basis for the set of all estimable functions.
   Solution:
   A basis for the set of estimable functions is

$$\{\mu + \alpha_1 + \gamma_1 + \delta_1, \alpha_2 - \alpha_1, \gamma_2 - \gamma_1 + \delta_2 - \delta_1\}.$$

   (c) What is the rank of the model matrix X?
   Solution: By part(b), rank = 3

6. T23. Suppose that y follows a Gauss–Markov model in which

$$X = \begin{pmatrix} 1 & 1 & t \\ 1 & 2 & 2t \\ \vdots & \vdots & \vdots \\ 1 & t & t^2 \\ 1 & 1 & s \\ 1 & 2 & 2s \\ \vdots & \vdots & \vdots \\ 1 & s & s^2 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

Here, t and s are integers such that $t \geq 2$ and $s \geq 2$. [Note: The corresponding model is called the "conditional linear model" and has been used for data that are informatively right-censored; see, e.g., Wu and Bailey (1989).]

(a) Determine which of the following linear functions of the elements of β are always estimable, with no further assumptions on t and s.

(i) β_1

(ii) $\beta_1 + \beta_2 t + \beta_3 s$

(iii) $\beta_2 + \beta_3(s+t)/2$

Solution:

(i) We have $\beta_1 = (1,0,0)\beta = [2(1,1,t) - (1,2,2t)]\beta$. Therefore, β_1 is always estimable.

(ii) Let $d^T = (a^T, b^T) = (a_1, \ldots, a_t, b_1, \ldots, b_s)$. Then,

$$\begin{aligned} d^T X &= \left(\sum_{i=1}^{t} a_i + \sum_{j=1}^{s} b_j, \sum_{i=1}^{t} a_i i + \sum_{j=1}^{s} b_j j, t \sum_{i=1}^{t} a_i i + s \sum_{j=1}^{s} b_j j \right) \\ &\equiv (A_1, A_2 + A_3, tA_2 + sA_3). \end{aligned}$$

Thus, $\beta_1 + \beta_2 t + \beta_3 s = c^T \beta$ Where $c^T = (1, t, s)$. For this function to be estimable, there must exist A_1, A_2, A_3 such that $A_1 = 1$, $A_2 + A_3 = t$, and $tA_2 + sA_3 = s$. The last two equalities cannot simultaneously hold if $t = s \neq 1$. Therefore, $\beta_1 + \beta_2 t + \beta_3 s$ is not always estimable.

(iii)

$$\begin{aligned} \beta_2 + \frac{\beta_3(s+t)}{2} &= (0, 1, \frac{s+t}{2})\beta \\ &= \frac{1}{2}[(1,2,2t) - (1,1,t)] + \frac{1}{2}[(1,2,2s) - (1,1,s)]\beta \end{aligned}$$

Hence, it is always estimable.

(b) Determine necessary and sufficient conditions on t and s for all linear functions of the elements of β to be estimable.

Solution:

A necessary and sufficient condition is $s \neq t$. For sufficiency, assume $s \neq t$. From part (a), we know $(1,0,0) \in R(X)$. This means $(1,1,t) - (1,0,0) = (0,1,t) \in R(X)$ and $(1,1,s) - (1,0,0) = (0,1,s) \in R(X)$. Hence,

$$\frac{(0,1,t) - (0,1,s)}{t - s} = (0,0,1) \in R(X)$$

and

$$\frac{(1/t)(0,1,t)-(1/s)(0,1,s)}{(1/t)-(1/s)}=(0,1,0)\in R(X).$$

Thus, $R(X)=R^3$ and all linear functions of β are estimable.

For necessity, if $s=t$, then $\{(1,\ 0,\ 0),\ (0,\ 1,t)\}$ forms a basis for $R(X)$, implying that the dimensionality of $R(X)$ is 2, meaning not all linear functions of β are estimable.