Statistica Sinica: Supplement

AN ALGEBRA FOR THE CONDITIONAL MAIN EFFECT PARAMETERIZATION

Arman Sabbaghi

Purdue University

Supplementary Materials

S1. Proofs

In this section, we provide the proofs for the results in Sections 4, 5, and 6.

Proof of Proposition 2. We recognize that

$$\begin{aligned}\left\langle X^{s}_{i|j}, X^{s'}_{l|k} \mid \mathcal{F}\right\rangle &= 2^{-2}\left\langle X_{\{i\}} + sX_{\{i,j\}}, X_{\{l\}} + s'X_{\{l,k\}} \mid \mathcal{F}\right\rangle \\ &= 2^{-2}(\left\langle X_{\{i\}}, X_{\{l\}} \mid \mathcal{F}\right\rangle + s'\left\langle X_{\{i\}}, X_{\{l,k\}} \mid \mathcal{F}\right\rangle + s\left\langle X_{\{i,j\}}, X_{\{l\}} \mid \mathcal{F}\right\rangle \\ &\qquad + ss'\left\langle X_{\{i,j\}}, X_{\{l,k\}} \mid \mathcal{F}\right\rangle),\end{aligned}$$

and then apply Lemma 1 to each term in this expression. □

Proof of Proposition 3. We have that

$$\begin{aligned}\left\langle X^{s}_{i|j}, X_I \mid \mathcal{F}\right\rangle &= 2^{-1}\left\langle X_{\{i\}} + sX_{\{i,j\}}, X_I \mid \mathcal{F}\right\rangle \\ &= 2^{-1}(\left\langle X_{\{i\}}, X_I \mid \mathcal{F}\right\rangle + s\left\langle X_{\{i,j\}}, X_I \mid \mathcal{F}\right\rangle),\end{aligned}$$

and then apply Lemma 1 to each term in this expression. □

Proof of Corollary 3. Twin CMEs correspond to the functions $X^{+}_{i|j}$ and $X^{-}_{i|j}$ for distinct $i, j \in \{1, \ldots, r\}$, which are orthogonal by Proposition 2. □

Proof of Corollary 4. This property follows from Proposition 3, and also corresponds to Corollary 2. □

Proof of Corollary 5. Sibling CMEs correspond to the functions $X_{i|j}^{s}$ and $X_{i|k}^{s'}$ for distinct $i, j, k \in \{1, \ldots, r\}$ and any $s, s' \in \{-, +\}$. From Proposition 2, their inner product is $2^{-2} b_{\mathcal{F},\phi}$, and from Corollary 1, their correlation is 1/2. □

Proof of Corollary 6. Let CME($A_i \mid A_j s$) and CME($A_l \mid A_k s'$) denote non-twin CMEs in a family of a resolution IV fraction $\mathcal{F} \subseteq \mathcal{D}_r$, where $i, j, l, k \in \{1, \ldots, r\}$ with $i \neq j$, $l \neq k$, and $s, s' \in \{-, +\}$. In this case, INT(A_i, A_j) and INT(A_l, A_k) are fully aliased in $\mathcal{F}$. Then as $\mathcal{F}$ is resolution IV, $\left\langle X_{i|j}^{s}, X_{l|k}^{s'} \mid \mathcal{F} \right\rangle = 2^{-2} s s' b_{\mathcal{F},\phi}$ from Proposition 2, so that the non-twin CMEs are aliased.

Now let CME($A_i \mid A_j s$) and CME($A_l \mid A_k s'$) denote CMEs with different parents and non-aliased corresponding two-factor interactions. Then $\left\langle X_{i|j}^{s}, X_{l|k}^{s'} \mid \mathcal{F} \right\rangle = 0$ from Proposition 2, so that the CMEs are orthogonal. □

Proof of Corollary 7. An uncle-nephew effect pair corresponds to the functions $X_{i|j}^{s}$ and $X_{\{j\}}$ for distinct $i, j \in \{1, \ldots, r\}$ and any $s \in \{-, +\}$. As all indicator function coefficients involving one or two factors are zero in a regular design of resolution at least III, we obtain the result from Proposition 3. □

Proof of Corollary 8. Cousin CMEs correspond to the functions $X_{i|j}^{s}$ and $X_{l|j}^{s'}$ for distinct $i, j, l \in \{1, \ldots, r\}$ and any $s, s' \in \{-, +\}$. As all indicator function coefficients involving one, two, or three factors are zero in a regular design of resolution at least IV, we obtain this result from Proposition 2. □

Proof of Proposition 4. For distinct $i, j \in \{1, \ldots, r\}$ in a 2^{r-p}_{IV} design $\mathcal{F} \subseteq \mathcal{D}_r$, $\mathrm{CME}(A_i \mid A_j+)$ is clear if and only if $\mathrm{CME}(A_i \mid A_j-)$ is clear, and $\mathrm{CME}(A_i \mid A_j+)$ is clear if and only if $\mathrm{INT}(A_i, A_j)$ is clear. The first direction in the latter result follows from Corollary 2. For the second direction, note from Corollary 2 that, for any $I \in \mathcal{P}_r - \{\{i\}, \{i,j\}\}$ such that $|I| \in \{1, 2\}$, and for any $l \in \{1, \ldots, r\} - \{i, j\}$, $\{i\} \triangle I$ and $\{i, j\} \triangle \{l\}$ are not defining words. The second direction then follows by recognizing that $\mathrm{INT}(A_i, A_j)$ is clear if and only if for any $I \in \mathcal{P}_r - \{\{i, j\}\}$ such that $|I| = 2$, $\{i, j\} \triangle I$ is not a defining word. Therefore, for any two 2^{r-p}_{IV} designs $\mathcal{F}$ and $\mathcal{F}'$, $\mathcal{F}$ has more clear two-factor interactions than $\mathcal{F}'$ if and only if $\mathcal{F}$ has more clear CMEs than $\mathcal{F}'$. □

Proof of Lemma 2. This follows by translating the result of Cheng, Steinberg, and Sun (1999) and Cheng (2014, p. 172) on the connection between a regular design's count of defining words of length four and the numbers of two-factor interactions in its aliasing sets into CME terminology. Specifically, the number of pairs of aliased two-factor interactions that correspond to the distinct factor pairs among the CMEs in a family $t \in \{1, \ldots, T_{\mathcal{F}}\}$ of $\mathcal{F}$ is $N_t(\mathcal{F})\{N_t(\mathcal{F})-1\}/2$. Each such pair of aliased two-factor interactions corresponds to a defining word of length four in $\mathcal{F}$. As a defining word of length four gives rise to three distinct pairs of aliased two-factor interactions, we have that the number of defining words of length four is $\sum_{t=1}^{T_{\mathcal{F}}} N_t(\mathcal{F})\{N_t(\mathcal{F})-1\}/6$. A minimum aberration 2^{r-p}_{IV} design minimizes the number of defining words of length four, and hence minimizes $\sum_{t=1}^{T_{\mathcal{F}}} N_t(\mathcal{F})\{N_t(\mathcal{F})-1\}$ among 2_{IV}^{r-p} designs $\mathcal{F}$. □

Proof of Proposition 5. From Corollary 1, the absolute correlation between any two non-sibling CMEs that do not involve both of the same factors is either 0 or 1/2. Lemma 2 then yields the result. □

Proof of Lemma 3. For $c, d \in \{1, \ldots, q\}$, entry (c, d) of $M^{\mathsf{T}}M$ is the inner product of columns c and d of M. This inner product of column vectors is equivalent to the inner product of $Z_{(c)}$ and $Z_{(d)}$ using the indicator function of $\mathcal{F}$ in Definition 8, with a multiplicative factor of 2^r being introduced because the inner product of column vectors in M does not involve division by 2^r as in Definition 8. □

Proof of Proposition 6. The entries of $M^{\mathsf{T}}M$ are derived by combining Lemma 3 with Lemma 1 and Propositions 2 and 3 to calculate the inner products of pairs of traditional and conditional effects across the columns of M. The determinant of $M^{\mathsf{T}}M$ then follows by recognizing its block-matrix structure, and using the standard determinant formula for such matrices. □

S2. Conditional Main Effect Analysis of Practical Application

The data from the painted panel experiment of Lorenzen and Anderson (1993, p. 242–249), in the order of the execution of the runs, are contained in Tables 1 and 2. Estimates and tests for the factorial effects that correspond to the final ANOVA they performed are in Table 3. Note that in this table, complete aliasing of two-factor interactions are denoted by "=". The R^2 and adjusted R^2 vales for the corresponding linear regression model are approximately 0.95 and 0.91, respectively. Residual diagnostics in Figure 1 do not indicate either violations of the standard regression assumptions of independent and identically distributed Normal errors, or the existence of outliers. As we can see, ME(A_1), ME(A_2), ME(A_3), ME(A_4), ME(A_5), ME(A_8), INT(A_4, A_7), and INT(A_2, A_8) = INT(A_3, A_5) = INT(A_4, A_6) are statistically significant at the 0.05 level. We cannot conclude which of the latter three two-factor interactions are active by the traditional analysis because they are fully aliased in the design. To resolve this limitation, we use the CME analysis method of Su and Wu (2017, p. 5–6). We first note that, among the estimates of the main effects for the factors involved in these interactions, the estimate of ME(A_2) (-0.01875) is closest to the estimate of the aliased set of two-factor interactions (0.0175) in absolute value. Then as these two estimators have opposite signs, we conclude from Rule 1 in (Su and Wu, 2017, p. 5) that CME($A_2 \mid A_8-$) is significant. Our final set of significant effects under the CME analysis are thus ME(A_1), ME(A_2), ME(A_3), ME(A_4), ME(A_5), ME(A_8), INT(A_4, A_7), and CME($A_2 \mid A_8-$).

Table 1: Runs $1-16$ of the experiment of Lorenzen and Anderson (1993, p. 246).

Run	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	Film Build
1	+	+	−	−	−	−	+	+	0.15
2	−	+	−	+	+	−	−	−	0.16
3	+	−	−	−	−	−	−	−	0.19
4	−	−	+	−	−	+	+	+	0.38
5	+	+	+	+	+	+	+	+	0.22
6	+	−	+	−	+	−	+	−	0.35
7	−	+	+	−	+	−	+	+	0.30
8	−	−	+	+	+	+	+	−	0.26
9	−	+	+	−	−	+	−	−	0.30
10	−	−	−	−	+	+	−	+	0.15
11	+	−	+	+	−	−	+	+	0.27
12	+	+	−	+	+	−	+	−	0.16
13	+	−	+	+	+	+	−	−	0.28
14	+	+	−	+	−	+	−	+	0.16
15	−	+	+	+	+	+	−	+	0.27
16	+	+	+	+	−	−	−	−	0.27

Table 2: Runs 17 − 32 of the experiment of Lorenzen and Anderson (1993, p. 246).

Run	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	Film Build
17	−	+	+	+	−	−	+	−	0.28
18	−	−	−	+	−	+	−	−	0.22
19	−	+	−	+	−	+	+	+	0.20
20	+	+	+	−	−	+	+	−	0.32
21	+	−	−	+	+	−	−	+	0.13
22	−	−	−	+	+	−	+	+	0.16
23	−	−	−	−	−	−	+	−	0.34
24	−	−	+	−	+	−	−	−	0.35
25	+	−	+	−	−	+	−	+	0.28
26	−	+	−	−	+	+	+	−	0.21
27	−	+	−	−	−	−	−	+	0.20
28	+	+	+	−	+	−	−	+	0.27
29	+	−	−	−	+	+	+	+	0.14
30	−	−	+	+	−	−	−	+	0.27
31	+	−	−	+	−	+	+	−	0.15
32	+	+	−	−	+	+	−	−	0.15

Table 3: Summary of the estimates and tests for the factorial effects in the final ANOVA of Lorenzen and Anderson (1993, p. 248). Complete aliasing of any pair of two-factor interactions in the design is denoted by "=" here.

Factorial Effect	Estimate	F-Test Statistic	p-Value
$\mathrm{ME}(A_1)$	−0.035	20.035	3.8×10^{-4}
$\mathrm{ME}(A_2)$	−0.01875	5.75	2.9×10^{-2}
$\mathrm{ME}(A_3)$	0.1125	206.96	1.4×10^{-10}
$\mathrm{ME}(A_4)$	−0.03875	24.55	1.4×10^{-4}
$\mathrm{ME}(A_5)$	−0.02625	11.27	4.0×10^{-3}
$\mathrm{ME}(A_6)$	−0.01	1.64	2.2×10^{-1}
$\mathrm{ME}(A_7)$	0.015	3.68	7.3×10^{-2}
$\mathrm{ME}(A_8)$	−0.0275	12.37	2.9×10^{-3}
$\mathrm{INT}(A_1, A_3)$	0.01625	4.32	5.4×10^{-2}
$\mathrm{INT}(A_1, A_5)$	0.015	3.68	7.3×10^{-2}
$\mathrm{INT}(A_2, A_4) = \mathrm{INT}(A_6, A_8)$	0.01625	4.32	5.4×10^{-2}
$\mathrm{INT}(A_2, A_6) = \mathrm{INT}(A_4, A_8)$	0.015	3.68	7.3×10^{-2}
$\mathrm{INT}(A_2, A_8) = \mathrm{INT}(A_3, A_5) = \mathrm{INT}(A_4, A_6)$	0.0175	5.01	4.0×10^{-2}
$\mathrm{INT}(A_3, A_4) = \mathrm{INT}(A_5, A_6)$	−0.015	3.68	7.3×10^{-2}
$\mathrm{INT}(A_4, A_7)$	−0.0225	8.28	1.1×10^{-2}

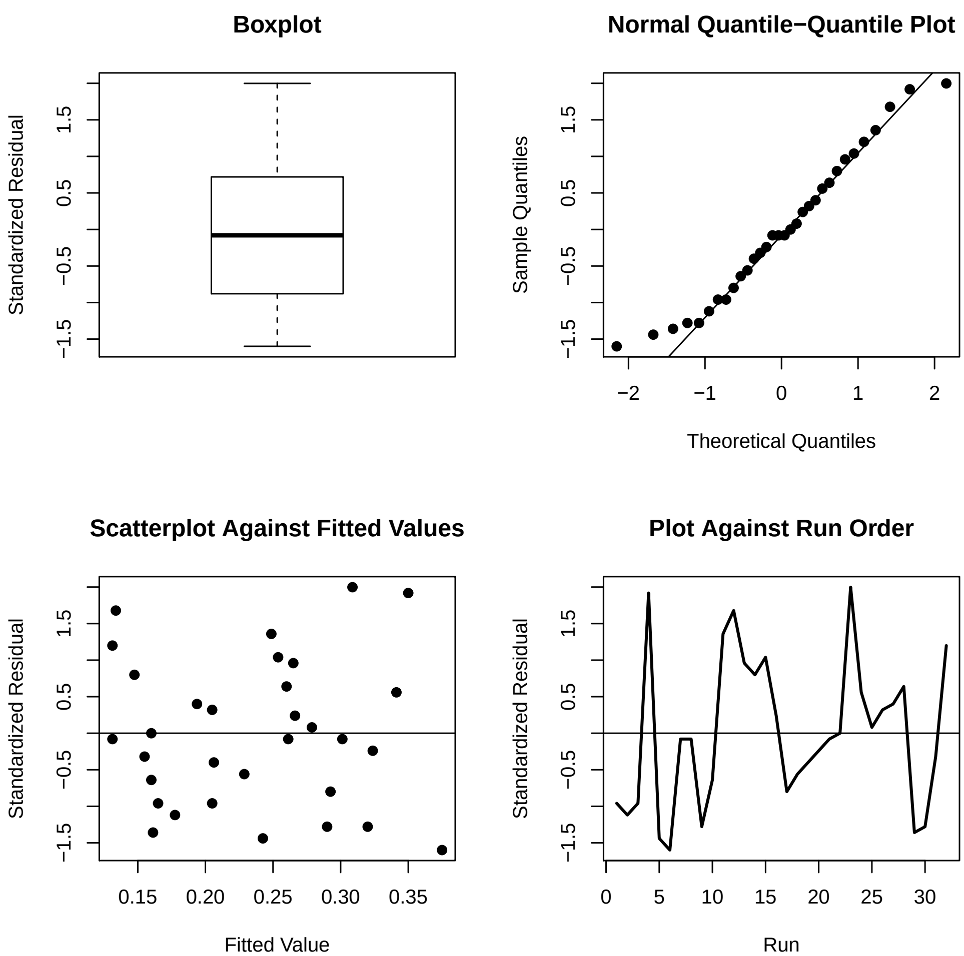

Figure 1: The boxplot, Normal quantile-quantile plot, scatterplot against fitted values, and plot against the run order, of the standardized residuals for the linear regression model that corresponds to the final ANOVA of Lorenzen and Anderson (1993, p. 242–249).

References

Cheng, C.-S. (2014). *Theory of Factorial Design: Single- and Multi-Stratum Experiments* (1 ed.). Monographs on Statistics and Applied Probability. CRC Press.

Cheng, C.-S., D. M. Steinberg, and D. Sun (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *J. R. Statist. Soc. Ser. B 61*(1), 85–93.

Lorenzen, T. J. and V. L. Anderson (1993). *Design of Experiments: A No-Name Approach* (1 ed.). Statistics: Textbooks and Monographs. CRC Press.

Su, H. and C. F. J. Wu (2017). CME Analysis: a new method for unraveling aliased effects in two-level fractional factorial experiments. *Journal of Quality Technology 49*(1), 1–10.

Department of Statistics, Purdue University, West Lafayette, IN 47907, USA.

E-mail: sabbaghi@purdue.edu

Phone: 765-496-0234

Fax: 765-494-0558