# Orthogonal Main-Effect Plans through Collapsing

- Method of collapsing factors: useful for constructing plans that are not orthogonal arrays but are smaller in run size than what would be required by the use of orthogonal arrays.

For an OA, need at least 18 runs

Table 6: Construction of $OME(9, 2^1 3^3)$ from $OA(9, 3^4)$ through Collapsing Factor

OA

OME NOA

| Run | A | B | C | D |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 2 | 1 |
| 3 | 0 | 2 | 1 | 2 |
| 4 | 1 | 0 | 1 | 1 |
| 5 | 1 | 1 | 0 | 2 |
| 6 | 1 | 2 | 2 | 0 |
| 7 | 2 | 0 | 2 | 2 |
| 8 | 2 | 1 | 1 | 0 |
| 9 | 2 | 2 | 0 | 1 |

$\longrightarrow$

| Run | A′ | B | C | D |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 2 | 1 |
| 3 | 0 | 2 | 1 | 2 |
| 4 | 1 | 0 | 1 | 1 |
| 5 | 1 | 1 | 0 | 2 |
| 6 | 1 | 2 | 2 | 0 |
| 7 | 1 | 0 | 2 | 2 |
| 8 | 1 | 1 | 1 | 0 |
| 9 | 1 | 2 | 0 | 1 |

nearly OA

alternative.

not balanced



# Orthogonal Main-Effect Plans through Collapsing

can check orthogonality for any two spaces

- Orthogonal main-effect ($OME$) plan: a plan in which all the main effect estimates are orthogonal (i.e., uncorrelated).

$\mathbb{1} \perp \Omega_A \perp \Omega_B \perp \cdots \perp \Omega_H$

- Proportional frequencies

$\mathbb{1}$ & all MEs $\rightarrow$ model matrix $\rightarrow$ $[\,\mathbb{1} \mid A \mid B \mid \cdots \mid H\,]$

  - For two factors in the design matrix, denoted by $A$ and $B$, let $r$ and $s$ be the number of levels of $A$ and $B$, respectively.
  - Let $n_{ij}$ be the number of observations at level $i$ of $A$ and level $j$ of $B$ in the two-factor projected design. $(A,B)=(i,j)$
  - We say that $n_{ij}$ are in *proportional frequencies* if they satisfy

$$n_{ij} = \frac{n_{i\cdot}n_{\cdot j}}{n_{\cdot\cdot}}, \Rightarrow \frac{n_{ij}}{n_{\cdot\cdot}} = \frac{n_{i\cdot}}{n_{\cdot\cdot}} \times \frac{n_{\cdot j}}{n_{\cdot\cdot}}$$

(*)

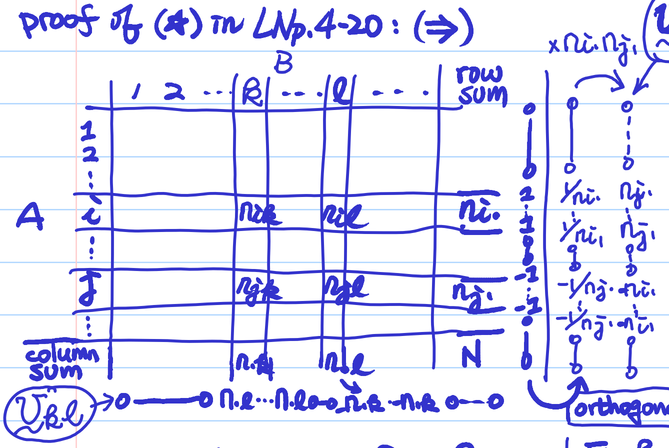
where $n_{i\cdot} = \sum_{j=1}^{s} n_{ij}$, $n_{\cdot j} = \sum_{i=1}^{r} n_{ij}$ and $n_{\cdot\cdot} = \sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij}$.

- In the main-effect-only model, the main effect estimates of factor $A$ and of factor $B$ are uncorrelated (orthogonal) if and only if $n_{ij}$ are in proportional frequencies.

# Orthogonal Main-Effect Plans through Collapsing

Table 7: Factors $A'$ and $B$ in $OME(9, 2^1 3^3)$ appear in proportional frequencies, where $n_{ij}$ appear as cell counts with $n_{i\cdot}$ and $n_{\cdot j}$ as margin totals

proof of (#) in LN p.4-20: (⇒)




| $A'$ \ $B$ ($s=3$) | 0 | 1 | 2 | Row Sum ($n_{i\cdot}$) |
|---|---|---|---|---|
| 0 ($r=2$) | 1 ($1/9$) | 1 | 1 | 3 ($1/3$) |
| 1 | 2 | 2 ($2/9$) | 2 | 6 ($2/3$) |
| Column Sum ($n_{\cdot j}$) | 3 ($1/3$) | 3 ($1/3$) | 3 ($1/3$) | 9 |

∵ $v_{kl} \in \Omega_B$, $u_{ij} \in \Omega_A$, & $\Omega_A \perp \Omega_B$

∴ $0 = v_{kl}^T u_{ij}$
$= n_{ik}(n_{\cdot l}\times n_{j\cdot}) + n_{jl}(n_{\cdot k}\times n_{i\cdot}) - n_{jk}(n_{\cdot l}\times n_{i\cdot}) - n_{il}(n_{\cdot k}\times n_{j\cdot})$

$\forall\ 1 \le i < j \le a \leftarrow$ # of levels of A

$\forall\ 1 \le k < l \le b \leftarrow$ : : : : B

(⇐) trivial.

$n_{\cdot k}(n_{i\cdot}n_{jl} - n_{j\cdot}n_{il}) = n_{\cdot l}(n_{i\cdot}n_{jk} - n_{j\cdot}n_{ik})$

Fix k, l, i, sum over j, including j = i, $\sum_{j=1}^{a}$

$n_{\cdot k}(n_{i\cdot}n_{\cdot l} - N\cdot n_{il}) = n_{\cdot l}(n_{i\cdot}n_{\cdot k} - N n_{ik})$

Fix k, i, sum over all l including l = k, $\sum_{l=1}^{b}$

$n_{\cdot k}(\underbrace{n_{i\cdot}N - N n_{i\cdot}}_{=0}) = N(n_{i\cdot}n_{\cdot k} - N n_{ik})$

$\Rightarrow n_{i\cdot}n_{\cdot k} = N\cdot n_{ik} \Rightarrow n_{ik} = n_{i\cdot}n_{\cdot k}/N.$



# Orthogonal Main-Effect Plans through Collapsing

- Construction of *OME* plans from *OA*s through collapsing
  - We can replace a factor with $t$ levels in an *OA* by a new factor with $s$ levels, $s < t$, by using a many-to-one correspondence between the $t$ levels and the $s$ levels.

Table 8: Collapsing a Four-Level Factor to a Three-Level Factor

$n_{ij} = n_{i\cdot}\times n_{\cdot j}/N$
$n_{ij'} = n_{i\cdot}\times n_{\cdot j'}/N$ ] $OME_{1}$
$n_{ij}+n_{ij'} = n_{i\cdot}(n_{\cdot j}+n_{\cdot j'})/N \leftarrow OME_2$

| $t=4$ | → $s=3$ | or | → $s=2$ | |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 1 | 0 |
| 2 | 1 | 2 | 1 | 0 |
| 3 | 2 | 2 | 0 | 1 |

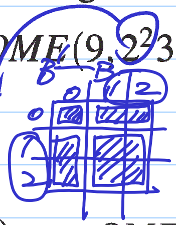
better (∵ balanced)

  - Because the cell frequencies between any two factors of an *OA* are equal, the collapsed factor and any remaining factor in the orthogonal array have proportional frequencies.
  - The method of collapsing factors can be $\boxed{\text{repeatedly applied}}$ to different columns of an *OA* to generate a variety of *OME* plans. ($OA \to OME_1 \to OME_2 \to OME_3 \cdots$)

# Orthogonal Main-Effect Plans through Collapsing

- Construction of *OME* plans from *OA*s through collapsing - Examples
  - $OA(9,3^4) \longrightarrow OME(9,2^1 3^3) \longrightarrow OME(9,2^2 3^2) \longrightarrow OME(9,2^3 3^1) \longrightarrow OME(9,2^4)$

    12 for OA    A'    A    1:2:2:4    more serious imbalance



  - $OA(8,4^1 2^4) \longrightarrow OME(8,3^1 2^4)$
  - $OA(16,4^2 2^9) \longrightarrow OME(16,4^1 3^1 2^9) \longrightarrow OME(16,3^2 2^9)$

    36 for OA

- Advantage of *OME* plans over *OA*s: run size economy
- Disadvantage of *OME* plans comparied to *OA*s

  e.g. → 3 levels $\xrightarrow{\text{collapsing}}$ 2 levels
  Why assign 2/3 N → level 0, 1/3 N → level 1,
  estimation of its mean is more accurate

  - Imbalance of run size allocation between levels
  - Loss in estimation efficiency
    * But, the estimation efficiency of *OME* plans is generally very high unless the number of levels of the collapsed factor is much smaller than that of the original factor.



# Orthogonal Main-Effect Plans through Collapsing

- Estimation efficiency comparison between *OME* plans and *OA*s
  - Collapsing a 3-level factor to a 2-level factor: assume that the total run size is $3m$ and the run size allocation between the 2 levels is $2m$ and $m$
    * Variance of its main effect estimate is

      $\sum y_i/2m$    $\sum y_i/m$    $\sigma^2\left(\frac{2}{3m}+\frac{2}{3m}\right)=$

      Var$(\bar{y}_0+\bar{y}_2-2\bar{y}_1)$    OME imbalanced

      $$Var(\bar{y}_1-\bar{y}_0)=\sigma^2(1/m+1/2m)=(3/2m)\sigma^2.$$

      $\sigma^2\left(\frac{1}{m}+\frac{1}{m}+4\frac{1}{2m}\right)$
      $\sigma^2\left(\frac{3}{4m}+\frac{3}{4m}+4\frac{3}{4m}\right)$

    * For balanced allocation of $\frac{3m}{2}$ runs to each levels, the variance is $\frac{4}{3m}\sigma^2$
    * Relative efficiency: $\frac{8}{9}$

      Var$(\bar{y}_2-\bar{y}_0)$ $\begin{cases}\sigma^2\left(\frac{1}{m}+\frac{1}{m}\right)\\ \sigma^2\left(\frac{3}{4m}+\frac{3}{4m}\right)\end{cases}$

      4m

  - Collapsing a 4-level factor to a 3-level factor: relative efficiency for linear main effect is $\frac{3}{4}$, and relative efficiency for quadratic main effect is $\frac{9}{8}$ if the middle level receives one-half of the runs

    $>1$

    | 0 | 1 | 2 | 0 | 1 | 2 |
    |---|---|---|---|---|---|
    | m | 2m | m | 4m/3 | 4m/3 | 4m/3 |

- *OME* plans constructed through collapsing fewer factors are preferred over those through collapsing more factors because the collapsing of each factor would result in the loss of degrees of freedom and of estimation efficiency. Example: $OME(8,3^1 2^4)$ and the $OME(9,3^1 2^3)$

❖ **Reading**: textbook, 8.8

# Complex Aliasing

$X=[X_1\ X_2]$, $\beta=\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}$, $Y = X\beta+\varepsilon \rightarrow \begin{pmatrix}\hat\beta_1^*\\ \hat\beta_2^*\end{pmatrix} = (X^TX)^{-1}X^TY \Rightarrow E(\hat\beta_1^*)=\beta_1$, $E(\hat\beta_2^*)=\beta_2$

$12\Big(\big[\,|\cdots|\,|\,|\,|\,|\,|\,\big]\Big)$ span a 12-dim space

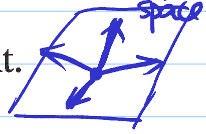
- Alias matrix
  - True model: $\mathbf{Y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon$, with $E(\varepsilon) = \mathbf{0}$ and $Var(\varepsilon) = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. may be an unidentifiable model.
  - Working model: $\mathbf{Y} = \mathbf{X}_1\beta_1 + \varepsilon$, where $\mathbf{X}_1$ is nonsingular. under-fitting.
  - The ordinary least squares estimator of $\beta_1$ under the working model is
  $$\hat\beta_1 = (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{Y}$$
  $\rightarrow E(\hat\beta_1)=\beta_1$ under working model.
  - Under the true model,
  ① $\beta_2\approx 0$ ② $X_1^TX_2\approx \mathbf{0}$ ③ if $X_1^TX_2 \not\approx \mathbf{0}$, $\beta_2\not\approx 0$
  $$\begin{aligned} E(\hat\beta_1) &= (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\,E(\mathbf{Y})\\ &= (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T(\mathbf{X}_1\beta_1+\mathbf{X}_2\beta_2)\\ &= \beta_1 + (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2\beta_2\end{aligned}$$
  $E(\hat\beta_{1,i}) = \beta_{1,i} + L_i^T\beta_2$

  where $(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2\beta_2$ is the bias of $\hat\beta_1$ under the true model.
  - We call $L = (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2$ the **alias matrix** because it provides the aliasing coefficients for the estimate $\hat\beta_1$.



# Complex Aliasing

- Alias matrix - example




  - Consider the $OA(12,2^7)$ used in the cast fatigue experiment.
  - Suppose that $\mathbf{X}_1$ consists of the 7 main effects. ← main effect only model
  - Suppose that $\mathbf{X}_2$ consists of the $\binom{7}{2} = 21$ two-factor interactions.
  - The true model is unidentiable. $(\because 1+7+21 > 12)$
  - Under the working model, the main effects estimates are uncorrelated and unbiased. The main effect estimate (say, of $A$) is
  $$\hat A = \bar y(A=+) - \bar y(A=-)$$
  the 29 effects span a 12-dim space, & the 29 effects are different ← different vectors

  But, if use $OA(12,2^4)$, $X_1'$: 4 ME's, $X_2'$: 6 2fi's, can fit $X_1'\beta_1 + X_2'\beta_2$ simultaneously & $E(\hat\beta_1^*)=\beta_1$.
  - Under the true model,
  $$\begin{aligned} E(\hat A) = A &-\tfrac{1}{3}BC - \tfrac{1}{3}BD - \tfrac{1}{3}BE + \tfrac{1}{3}BF - \tfrac{1}{3}BG\\ &+\tfrac{1}{3}CD - \tfrac{1}{3}CE - \tfrac{1}{3}CF + \tfrac{1}{3}CG + \tfrac{1}{3}DE\\ &+\tfrac{1}{3}DF - \tfrac{1}{3}DG - \tfrac{1}{3}EF - \tfrac{1}{3}EG - \tfrac{1}{3}FG\end{aligned}$$
  $L = (X_1^TX_1)^{-1}(X_1^TX_2) = \begin{pmatrix}1/12 & & 0\\ & \ddots & \\ 0 & & 1/12\end{pmatrix} X_1^TX_2 = \big(\ \mathrm{cor}(ME, 2fi)\ \big)$

  same if do $OA(12,2^{11})$
  - For the 12-run design, each main effect has the same property as that for factor $A$ in that it is partially aliased with the 2-factor interactions not involving the main effect and the aliasing coefficients are $1/3$ or $-1/3$.