**Design for Order-of-Addition Experiments**

*Dennis Lin*
University Distinguished Professor  
Department of Statistics  
The Pennsylvania State University

*at*
Department of Statistics, Temple University  
4 October, 2017

---

**Design of Experiment**

*How to collect Useful Information?*

---

**Design of Experiment**

*Analyzing historical data is like listening to a lecture  
Running a designed experiment is like conducting an interview*

---

**Randomization in Theory vs. Randomization in Reality**

*R.A. Fisher (1920)*
How Should the Data be collected?

Randomly
or
Systematically

Some Lin recent design projects
- Computer Experiment—LHC & UD
- Order of Addition Experiment
- Run Order Consideration
- t-covering array
- Design for On-Line Experiment
- New Type of Composite Design
- Fake Factors for estimate \( \sigma^2 \)
- Meta-Analysis

Lady and Tea Tasting (Fisher)

Lady and Tea Tasting (Fisher)

Three-Cup Chicken

Three Cups:
- Soy Sauce
- Wine
- Sesame Oil

Which first? Which last? Does it matter?

Three-Cup Chicken

Three Cups:
- Soy Sauce
- Wine
- Sesame Oil

Which first? Which last? Does it matter?
There are $m!$ possible combinations, how could we run fraction of them?

For three components, there are $3! = 6$ possible “treatments” to be tested.

In general, there are $m!$ treatments to be tested.

for example, $10! = 3,638,800$. This may not be feasible.

Order of addition (OofA) experiment:

the requirement for Ran, Crm1 and NXT1, etc

Journal of Cell Biology (2001)

—$m$ is about 10.

OofA in Genetics Areas

The construction of phylogenetic trees depends on the order of taxa

Many taxa (more than 10) are involved...

Often, a set of random orders are tested (Olsen et al. 1994, Stewart et al. 2001)

How to choose a subset of the orders? Randomly or systematically???
OofA in Different Areas

- Food science: Fuleki and Francis (1968)
- Food science: Jourdain et al. (2009)
- Nutritional science: Karim et al. (2000)
- Pharmaceutical science: Rajaonarivony et al. (1993)

Experiments are needed to find the optimal addition order!

Research Issues

How to run (small) \( n \), among those \( m! \) experiments, to find out the “optimal” sequence/order-of-addition (OofA)?

Note: \( 10! = 3,628,800 \)

Order-of-Addition Experiment

Linking to conventional design...

- What are the experimental variables (X,’s)?
- What is the experimental unit?

Outline

- Introduction (baby optimal design)
- Model Formulation (PWO)
- Optimality of the Full PWO Design
- Orthogonality of a PWO Design
- Minimal-point PWO Design
- Optimal Fractional PWO Design
- Conclusion and Future Work
**Brief on Optimal Design**

**Matrix Form**

\[ y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \quad i = 1, 2, \ldots, n \]

\[
\begin{bmatrix}
 y_1 \\
 y_2 \\
 \vdots \\
 y_n \\
\end{bmatrix} = 
\begin{bmatrix}
 1 & x_{11} & x_{21} & \cdots & x_{k1} \\
 1 & x_{12} & x_{22} & \cdots & x_{k2} \\
 \vdots & \vdots & \vdots & \cdots & \vdots \\
 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \\
\end{bmatrix}
\begin{bmatrix}
 \beta_0 \\
 \beta_1 \\
 \beta_2 \\
 \vdots \\
 \beta_k \\
\end{bmatrix}
+ 
\begin{bmatrix}
 \varepsilon_1 \\
 \varepsilon_2 \\
 \vdots \\
 \varepsilon_n \\
\end{bmatrix}
\]

or \( Y = X\beta + \varepsilon \)

**Estimation**

\[ \hat{\beta} = (X'X)^{-1}X'Y \]

LSE / MLE under i.i.d. Normal

\[ \text{Var}(\hat{\beta}) = (X'X)^{-1} \cdot \sigma^2 \]

Assuming (say) \( \varepsilon_i \sim N(0, \sigma^2) \) i.i.d.

**Design Issue:**

*Now, suppose you have full control on the X matrix…*

Choose \( X \) such that \((X'X)^{-1}\) is minimized—or \( X'X \) is maximized (in some senses).

**Optimal Design—General Setting**

- Given the model \( y = f(x) + \varepsilon \),
- find its information matrix \( \mathbf{I} \),
  - The Optimal Design \( X \) is the design which “maximizes” the information matrix \( \mathbf{I} \).
- For Linear Model \( y = X\beta + \varepsilon \),
  - the information matrix is \( \mathbf{I} = X'X \).
Optimality Theorem

the full PWO design is optimal under:

- **D-criterion** = \( \arg \max \det(M)^{1/p} \), \( p = \left( \frac{m}{2} \right) + 1 \)
- **A-criterion** = \( \arg \min \text{tr} \left( M^{-1} \right) \)
- **E-criterion** = \( \arg \max \lambda_{\min}(M) \)
- **M.S.-criterion** = \( \arg \min \text{tr} \left( M^2 \right) \)

where, \( M = \text{Information Matrix} \)

Continuous Version

General estimator:

\[ \hat{\theta}_c = \int y(x)\zeta(dx), \]

where \( \zeta(dx) \) is a signed vector-measure.

\[ \hat{\theta}_{OLSE} = \int y(x)M^{-1}(\xi) f(x)\xi(dx), \]

where

\[ M(\xi) = \int f(x) r^T(x) \xi(dx), \]

and \( \xi(dx) \) is a design (probability measure for OLSE; a signed measure for SLSE). The covariance matrix of \( \hat{\theta}_{OLSE} \) is

\[ \text{Var} \left( \hat{\theta}_{OLSE} \right) = M(\xi)^{-1} \left[ \int \int K(x,z) f(x) r^T(z) \xi(dx) \xi'(dz) \right] M(\xi)^{-1} \]

Pairwise-order (PWO) model

Van Nostrand (1995)

- Suppose there are \( m \) components to be added, denoted by 1, 2, ..., \( m \)
- For any order \( \alpha \) and \( 1 \leq j < k \leq m \), define the PWO factor

\[ z_{jk}(\alpha) = \begin{cases} 
1 & \text{if } j \text{ precedes } k \text{ in } \alpha, \\
-1 & \text{if } k \text{ precedes } j \text{ in } \alpha.
\end{cases} \]

For example, \( \alpha = 312 \) implies

\[ z_{12} = +1, \ z_{13} = -1, \ \text{and} \ z_{23} = -1 \]
**Problem Formulation** *(m=3 example)*

<table>
<thead>
<tr>
<th>Sequence</th>
<th>I_{1-2}</th>
<th>I_{1-3}</th>
<th>I_{2-3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1 3 2</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>2 1 3</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2 3 1</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>3 1 2</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>3 2 1</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

**Model**

\[ y = M + \beta_{1,2}I_{1,2} + \beta_{1,3}I_{1,3} + \beta_{2,3}I_{2,3} + \epsilon \]

**Test**

\[ H_0: \beta_{i,j} = 0 \]

**PWO model**

*For any order* \( \alpha \) *affects the response via the Pairwise-order (PWO) effect*

\[ \tau(\alpha) = \beta_0 + \sum_{1 \leq j < k \leq m} z_{jk}(\alpha)\beta_{jk}, \]

\( \tau(\alpha) \): expected response arising from \( \alpha \)

\( \beta_{jk} \)'s: linear coefficients to estimate

*With* \( m \) *components, there are* \( \binom{m}{2} \) *PWO factors.*

**Research Issues**

How to run (small) \( n \) among \( m! \) many experiments to test all \( \binom{m}{2} \) pairwise order?

\[ H_0: \beta_{i,j} = 0 \]

**Full PWO Design**

*PWO design:* \( [z_{jk}(\alpha_i)]_{jk} \)

*Full PWO design* \( (Z_F) \): representing all the permutations

<table>
<thead>
<tr>
<th>Sequence</th>
<th>I_{1-2}</th>
<th>I_{1-3}</th>
<th>I_{2-3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1 3 2</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>2 1 3</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2 3 1</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>3 1 2</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>3 2 1</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

As compared with \( 2^3 \) Full Factorial design, the treatment \((+−+)\) and \((−+−)\) are not feasible.
The moment matrix (information matrix) of full PWO design:

$$M_f = X_f^T X_f / N, \text{ with } X_f = [1, Z_f] \text{ and } N = m!$$

for $m = 4$, $M_f = \text{diag}(1, \tilde{M}_f)$ and

$$\tilde{M}_f = \begin{bmatrix}
1 & 1/3 & 1/3 & -1/3 & -1/3 & 0 \\
1/3 & 1 & 1/3 & 1/3 & 0 & -1/3 \\
1/3 & 1/3 & 1 & 0 & 1/3 & -1/3 \\
-1/3 & 1/3 & 0 & 1 & 1/3 & 1/3 \\
-1/3 & 0 & 1/3 & 1/3 & 1 & 1/3 \\
0 & -1/3 & 1/3 & -1/3 & 1/3 & 1
\end{bmatrix}.$$
**Optimality Theorem**

The full PWO design is optimal under:

- **D-criterion** = $\arg \max \det(M)^{1/p}$, $p = \binom{m}{2} + 1$
- **A-criterion** = $\arg \min \tr(M^{-1})$
- **E-criterion** = $\arg \max \lambda_{\text{min}}(M)$
- **M.S.-criterion** = $\arg \min \tr(M^2)$

**Explicit Values for the Optimality Criteria**

Explicit values of the $D-/A-/E-/M.S.$-criteria are needed for comparative purpose.

- Benchmarks to assess the efficiency of any smaller design
- To derive such criteria, the eigen-structure of $M_f$ is investigated

**Theorem 2**

$M_f$ has eigenvalues $1$, $(m+1)/3$ and $1/3$, with multiplicities $1$, $m-1$ and $\binom{m-1}{2}$, respectively. Then for the full design:

- **D-criterion** = $\left[\det(M_f)\right]^{1/p} = \left(\frac{(m+1)^{m-1}}{3^m}\right)^{\frac{1}{p}}$
- **A-criterion** = $\tr(M_f^{-1}) = 1 + \frac{3m(m-1)^2}{2(m+1)}$
- **E-criterion** = $\lambda_{\text{min}}(M_f) = \frac{1}{3}$ and
- **M.S.-criterion** = $\tr(M_f^2) = 1 + \frac{(m-1)m(2m+5)}{18}$

$p = (m+1)/2 + 1$

**Optimality Theorem: A Catch!**

A (fractional) PWO design is $D-/A-/M.S.$-optimal

**If and only if**

it has the same moment matrix as the full PWO design.
**Constraint on the Correlation**

**Lemma 1**

For any (not full) PWO design with $m$ components, and any $1 \leq j < k < l \leq m$, it always holds that

$$\bar{M}(jk, jl) - \bar{M}(jk, kl) + \bar{M}(jl, kl) = 1,$$

Note: $\bar{M}(jk, jl)$ indicates the correlation between PWO factors $Z_{jk}$ and $Z_{jl}$.

**Optimality Criteria: Another Catch!!**

- PWO design can **NOT** be perfectly orthogonal — 
  - no regular fractional factorial design can be used as a PWO design.

- The maximum correlation ($r_{max}$) is at least $1/3$.

**A Bridge too far...**

Primitive Idea—

- **If** there exist an appropriate $2^{k-p}$ design for OofA experiment...

- Applying Cheng and Li (1993, *Technometrics*), choose the fraction to avoid those infeasible runs (due to transitive property).
  "Constructing Orthogonal Fractional Factorial Designs When Some Factor-Level Combinations Are Debarred"
Minimal-point PWO Designs

There are $m!$ possible runs, the minimal point PWO design requires $(\frac{m}{2}) + 1$ runs.

$m=3 \rightarrow m!=6 & (\frac{m}{2})+1=4$; which (best) 4 among those 6 runs?

$m=4 \rightarrow m!=24 & (\frac{m}{2})+1=7$; which (best) 7 among those 24 runs?
\[ m=3 \rightarrow m!=6 \text{ } \& \text{ } \binom{m}{2}+1=4; \]
which (best) 4 among those 6 runs?

There are \( \binom{6}{4} = 15 \) possibilities.

\[ m=4 \rightarrow m!=24 \text{ } \& \text{ } \binom{m}{2}+1=7; \]
which (best) 7 among those 24 runs?

There are \( \binom{24}{7} = 346,104 \) possibilities.

**Minimal-Point Design (m=3)**

Maximun D-efficiency: 0.71
Sample design: 123, 213, 132, 231

\[ \begin{array}{ccc}
123 & + & + & + \\
213 & - & + & + \\
132 & + & + & - \\
231 & - & - & + \\
\end{array} \]

D-efficiency of full design: 0.88
Relative Efficiency: 0.71/0.88 = 0.81

**Minimal-Point Design (m=4)**

Maximun D-efficiency: 0.70
Sample design: 1234, 2134, 1324, 3214, 1243, 2341, 3142

\[ \begin{array}{ccc}
1234 & + & + & + & + \\
2134 & - & + & + & + \\
1324 & + & + & - & + \\
3214 & - & - & + & + \\
1243 & + & + & + & + \\
2341 & - & - & + & - \\
3142 & + & - & - & - \\
\end{array} \]

D-efficiency of full design(4!=24): 0.78
Relative Efficiency: 0.70/0.78 = 0.90
For \( m \geq 5 \),

\[
\binom{m}{2} + 1 = \binom{120}{11} \text{ is too large!}
\]

need a systematic construction method!

**Construction of minimal-point OofA designs \((m \geq 6)\)**

- Take \( H_1 = (Q:1) \)
  - then \( H \) is a minimal-point OofA design.

\[
H = \begin{pmatrix}
H_1 & H_2 \\
H_3 & H_4
\end{pmatrix}
\]

\( H_2 \) is a matrix with all elements = \(-1\);
\( H_3 \) is a matrix with all elements = \(+1\);
\( H_4 = (h_{ij}) \) is a matrix with elements = \(+1\), if \( i \leq j \); and \(-1\) otherwise.

- its \( d \)-efficiency is

\[
D_d(H) = \left( \binom{\binom{n}{2} - n}{m} H[H_1] \right)^{1/(\binom{m}{2} + 1)}
\]
**D-efficiencies of the full PWO designs and the minimal-point designs**

<table>
<thead>
<tr>
<th>(m)</th>
<th>(m!)</th>
<th>(\binom{m}{2} + 1)</th>
<th>(D_e(H))</th>
<th>(D_e(F_{mn}))</th>
<th>(D_e(H))</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>6</td>
<td>4</td>
<td>0.707</td>
<td>0.577</td>
<td>0.810</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>7</td>
<td>0.697</td>
<td>0.777</td>
<td>0.897</td>
</tr>
<tr>
<td>5</td>
<td>120</td>
<td>11</td>
<td>0.591</td>
<td>0.706</td>
<td>0.837</td>
</tr>
<tr>
<td>6</td>
<td>720</td>
<td>16</td>
<td>0.349</td>
<td>0.656</td>
<td>0.532</td>
</tr>
<tr>
<td>7</td>
<td>5040</td>
<td>22</td>
<td>0.232</td>
<td>0.618</td>
<td>0.375</td>
</tr>
</tbody>
</table>

Note: \(R\) represents the minimal-point order of addition design; \(F_{mn}\) represents the pair-wise ordering design; \(D_e(H)\), \(D_e(F_{mn})\) and \(D_e(H)\) are the \(D\)-efficiencies of \(H\) and \(F_{mn}\), and the relative \(D\)-efficiency of \(H\), respectively.

---

**A Class of Optimal Fractional PWO Design**

**Information matrix of PWO Design**

The **moment matrix** (information matrix) of full PWO design:

\[
M_f = X_f^T X_f / N, \quad \text{with } X_f = [1, Z_f] \quad \text{and } N = m!
\]

For \(m = 4\), \(M_f = \text{diag}(1, \bar{M}_f)\) and

\[
\bar{M}_f = \begin{bmatrix}
1 & 1/3 & 1/3 & -1/3 & -1/3 & 0 \\
1/3 & 1 & 1/3 & 1/3 & 0 & -1/3 \\
1/3 & 1/3 & 1 & 0 & 1/3 & 1/3 \\
-1/3 & 1/3 & 0 & 1 & 1/3 & -1/3 \\
-1/3 & 0 & 1/3 & 1/3 & 1 & 1/3 \\
0 & -1/3 & 1/3 & -1/3 & 1/3 & 1
\end{bmatrix}
\]
An example of Optimal PWO Design

For \( m=4 \) \((m!)=24\), the following (half) fractional PWO design is “optimal.”

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
2 & 1 & 4 & 3 \\
4 & 3 & 1 & 2 \\
3 & 4 & 2 & 1 \\
1 & 3 & 2 & 4 \\
3 & 1 & 4 & 2 \\
4 & 2 & 1 & 3 \\
2 & 4 & 3 & 1 \\
1 & 4 & 2 & 3 \\
4 & 1 & 3 & 2 \\
3 & 2 & 1 & 4 \\
2 & 3 & 4 & 1 \\
\end{array}
\]

\( \overline{M}_f = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{1}{3} & 1 & \frac{1}{3} & \frac{1}{3} & 0 & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & 1 & 0 & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & 0 & 1 & \frac{1}{3} & 1 & \frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 1 \end{bmatrix} \)

It share the same moment matrix with the full PWO design.

An Example of Optimal Design

This design entails a partitioned structure:

\[
\begin{bmatrix}
B_1 & \bar{B}_1 \\
\bar{B}_1 & B_1 \\
B_2 & \bar{B}_2 \\
\bar{B}_2 & B_2 \\
B_3 & \bar{B}_3 \\
\bar{B}_3 & B_3 \\
\end{bmatrix}, \quad \text{with } B_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad \bar{B}_1 = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix}.
\]

\[
B_2 = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}, \quad \bar{B}_2 = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix},
\]

\[
B_3 = \begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix}, \quad \bar{B}_3 = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}.
\]

Extension to larger \( m \)

Such a construction method can be extended to any larger, even \( m \).

\textbf{Theorem 3.}

For any \( m \geq 4 \) and any \( 2 \leq r \leq m/2 \), there exist optimal PWO designs with \( m \) components and \( m!/r! \) runs.

Under Construction...

\textbf{A Class of Optimal Fractional PWO Design}

—an optimal PWD design with \( 2 \cdot \binom{m}{2} \) runs.
**Conclusion**
- The optimality theory for PWO designs
- The explicit values of optimality criteria
- Description on the orthogonality of any PWO design
- Systematic construction of efficient minimal-point PWO designs
- Systematic construction of optimal fractional PWO designs

**Order-of-Addition Experiments**
- Data Analysis Strategies
- Beyond PWO system
  - Triple-wise order (TWO) system?
  - Travel Salesman Problem (TSP)
  - etc
- Run order consideration

**Some (Very Selective) References**

**Send $1000 to**
- Dennis Lin
  - University Distinguished Professor
  - 317 Thomas Building
  - Department of Statistics
  - Penn State University
- +1 814 865-0377 (phone)
- +1 814 863-7114 (fax)
- DennisLin@psu.edu

(Customer Satisfaction or your money back!)