Evaluating and Constructing Designs for Robustness to Unusable Observations

Byran Smucker\textsuperscript{1}, Willis Jensen\textsuperscript{2}, Zichen Wu\textsuperscript{1}, and Bo Wang\textsuperscript{1}

\textsuperscript{1}Miami University, Oxford, OH
\textsuperscript{2}W.L. Gore and Associates, Flagstaff, AZ

March 2, 2017
IFPAC Annual Meeting, North Bethesda, MD
Study Finds 79% Of Statistics Now Sobering

NEWS IN BRIEF
August 20, 2014
VOL. 50 ISSUE 33
News
Optimality vs. Robustness

A Paradox?
Box and Draper (1975)

Addressed 14 ways a response surface design can be good.

- Generate information in region of interest
- Ensure fitted values are close as possible to true values
- Detect lack of fit
- Allow for transformations
- Allow for blocks
- Allows for building up of sequential experiments
- Provide internal estimate of variability
- Robust to wild observations and non-normality
- Uses minimum number of runs
- Allows for graphical assessment
- Simple to calculate
- Robust to the factors settings (the $x$’s)
- Do not require a lot of levels in the $x$’s
- Provide check of constant variance assumption
Robustness to Missing Observations

Note: We distinguish between robustness to outliers and robustness to missing observations.
How Many Have Ever Had an Experiment With Missing Observations?

- **Prevalence of Missing Observations**
  - Siddiqui (2011) suggested 1-10% of observations are wild
  - Co-author’s experience suggests that values of 0-20% are possible

- **Assume Observations are Missing at Random**
  - Does not always hold: could be that factor ranges chosen poorly
  - Assume the missing values are not a result of the factor levels

- **Why Not Just Redo the Missing Runs?**
  - Sometimes this is possible
  - Other times extremely costly or impossible to do
Generic Example of Modern Industrial Process

These characteristics lead to difficulty in redoing missing runs.
Characteristics of Modern Industrial Processes

- Increasing process complexity (more steps and more variables)
- Increasing equipment scales (more challenging to use equipment for experiments)
- Increasing supply chain complexity (raw materials coming from multiple suppliers at multiple places in the process)
- Increasing physical distances covered by process (different steps in different facilities and geographies)
What other characteristics make it more difficult to redo missing runs?
Regression

Standard regression model: \( y = X\beta + \epsilon \).

Variance-covariance matrix of the least squares estimators is
\[
\text{Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}
\]
The \( X^T X \) matrix is called the \textit{information matrix}.

We consider two types of models:

\[
f^T(x)\beta = \beta_0 + \sum_{i=1}^{k} \beta_i x_i
\]

\[
f^T(x)\beta = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2.
\]
Classical Designs

Fractional factorial designs and Central Composite designs

Optimal Designs

Choose design to have desirable statistical properties.

1. D-optimal design relates to variance-covariance matrix of least squares estimators:

   \[ \text{Assume } f \rightarrow \text{Expand to } X \rightarrow \text{Choose } \xi_n \text{ to maximize } |X^TX| \]

2. I-optimal designs minimize the average prediction variance of the design across design space.

3. Designs constructed to be robust to missing observations.
Use the determinant of the information matrix, $|X^T X|$. 

What if $m$ observations are missing?

$$D_F(i, m) = \left( \frac{|X_{n-m,i}^T X_{n-m,i}|}{|(X^*)^T_n (X^*)_n|} \right)^{1/p}, \quad i = 1, 2, \ldots, \binom{n}{m}$$

This metric gives a sense, in an absolute way, of how much information is being lost when $m$ runs are missing.
Measuring Design Quality for Response Surface Models

I-optimal designs because they seek to minimize the average prediction variance across the design space:

\[ I = \int_R f^T(x)(X^T X)^{-1} f(x) \, dx, \]

If \( m \) observations are missing?

\[ I_F(i, m) = \frac{l_i^*}{l_{n-m,i}}, \quad i = 1, 2, \ldots, \binom{n}{m}, \]
How to Assess Impact of Missing Runs

Examine how much information classical and optimal designs lose when runs go missing.

For instance, if 1 run is missing from an $n$-run design, we will compute the D-efficiency for each possible $(n - 1)$-run design and look at its distribution.

We’ll look at first- and second-order response surface models, various design sizes, and a small number of missing runs ($m = 0, 1, 2, 3$).
**First-order Models:** \( k = 5 \)

**Figure:** D-efficiencies for possible main effects designs, for \( k = 5 \) and \( n = \{8, 10, 12, 16\} \), according to the number of missing runs.
Second-order Models: \( k = 3 \)

### Second-order Models: \( k = 3 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>12</th>
<th>16</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-Efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of Missing Runs

Designs:
- CCD
- I-Optimal
- THA

**Motivation**

**Assessment and Scope**

**Results**

**Conclusions**

Smucker et al. Robustness of Designs to Missing Observations
Conclusions: Screening Designs

1. Missing observations have relatively little impact on first-order designs unless design is small.
2. Fractional and D-optimal designs are robust to missing observations.
3. If you are worried about losing a run or two, add a few runs at the outset.
Conclusions: Response Surface Designs

1. No evidence that optimal designs are less robust to missing observations than classical designs
2. Optimal designs are more efficient, and the efficiency holds up when a few observations are lost.
3. Missing-robust optimal designs are better when original number of runs is small
1. Optimal designs and classical designs have similar robustness properties, in terms of missing runs.

2. For severely resource-constrained experiments for which you are concerned about missing observations, either (1) add a few extra runs; or (2) consider using a missing-robust design.
Some Unanswered Questions

- What is the impact on spherical regions (e.g., 5 level CCD vs. optimal)?
- What is the impact on other types of designs (blocking, split-plots, mixtures, etc)?
- Are there better metrics to assess impact (ability to detect significant effects, width of intervals, bias and variance in predictions)?
- How could we obtain a robustness criteria based on I-optimality?
Satirical headline: Professor celebrates landmark publication that will be carefully read by two people (h/t Maria Weese)

Reference
Industrial Statistics Virtual Collaboratory

Community Blogs

- **Update and SPAIG Award**
  
  By: Dr. Byran Jay Smucker  one month ago
  
  It has been sometime since we've posted in this space, mostly because of a lack of readily available material to share. We remain interested in unsolved problems from industry or academia. If you have an industrial statistics problem and need a research...

- **Collaborative Research Report: Accelerated destructive degradation test data analysis**
  
  By: Dr. Byran Jay Smucker  5 months ago
  
  Last time, we announced an RSS feed. Now, thanks to Lara Harmon and the folks at the ASA, we have an RSS button on the ISVC landing page. As always, we are looking for material for this space. If it has to do with academic/industry collaboration....

community.amstat.org/isvc; contact: Byran Smucker (smuckerb@miamioh.edu)
Questions?
Extra Slide. Second-order Models: \( k = 5 \)