The Construction of Missing-Robust Experimental Designs and their Comparison to Classical and Optimal Designs

Byran Smucker[1], Willis Jensen[2], Zichen Wu[1], and Bo Wang[1]

[1]Miami University, Oxford, OH

[2]W.L. Gore and Associates, Flagstaff, AZ

October 12, 2017
The Design and Analysis of Experiments Conference, UCLA

Outline

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

15^{th} Way a Design Can Be Good

Robustness to Missing Observations

15^{th} Way a Design Can Be Good

Robustness to Missing Observations

(We distinguish between robustness to outliers and robustness to missing observations.)

The Problem

How Many Have Ever Had an Experiment With Missing Observations?

The Problem

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

The Problem

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

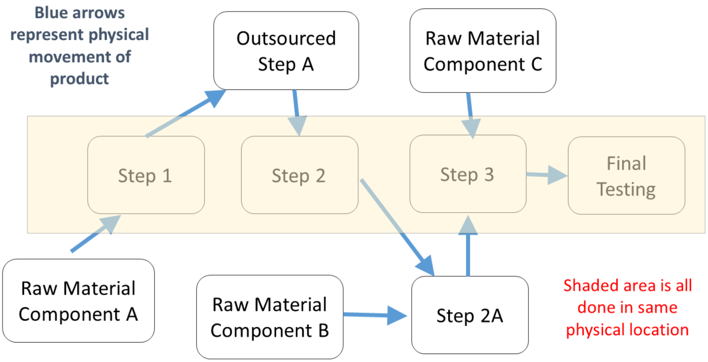
The Problem

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)

Goals for the Talk

1. Describe the construction of missing-robust designs
2. Compare these designs with classical and optimal designs in terms of their missing-robustness

Outline

Papers from a Pre-Computational Statistical Era

Some Considerations in the Optimal Design of Experiments in Non-optimal Situations

By AGNES M. HERZBERG and DAVID F. ANDREWS
Imperial College, London *University of Toronto*

[Received April 1976. Revised July 1976]

"*It is to be noted that when any part of this paper appears dull, there is a design in it.*"
SIR RICHARD STEELE, *The Tatler*, No. 38.

SUMMARY

Designs which satisfy optimal design criteria will usually not be optimal when some observations are missing, when outliers are present or when, contrary to the assumptions, the error of the observations arises from a non-Gaussian distribution. Designs robust against such short-comings will be introduced. They are slightly less than optimal under ideal conditions but will be more efficient under more realistic conditions. The probability of breakdown of a design is defined. Measures of robustness of a design are introduced in the cases of the generalized variance and the minimization of the maximum variance criteria. These measures and the probability of breakdown should be taken into account when the decision for a particular type of design is made. Some examples are given.

Keywords: EXPERIMENTAL DESIGN; ROBUSTNESS; OPTIMAL DESIGN; MISSING VALUES; OUTLIERS; NON-GAUSSIAN DISTRIBUTIONS

THE ROBUSTNESS AND OPTIMALITY OF RESPONSE SURFACE DESIGNS

David F. ANDREWS
University of Toronto

Agnes M. HERZBERG
Imperial College, London

Received 6 April 1977; revised manuscript received 16 December 1977
Recommended by V. Fedorov

Abstract: Recent work on extended optimality criteria for robust designs is applied to response surface problems. Methods of calculation are described and the criteria illustrated with several examples. The extended criteria discriminate among designs equivalent by other criteria.

Key words: Robustness; Optimal Design; Missing Values; Matrix Compounds.

From JRSS-B (1976) and JSPI (1979), respectively.

The Herzberg & Andrews (HA) Generalization of the D-criterion

Consider the following quantity: $\left|\mathbf{X}^T\mathbf{D}^2\mathbf{X}\right|^{\frac{1}{p}}$.

We have $\mathbf{D}^2$ a diagonal matrix with d_i^2 on the diagonal, where:

$$d_i^2 = \begin{cases} 0 & \text{with probability} & \alpha(\mathbf{x}), \\ 1 & \text{with probability} & 1-\alpha(\mathbf{x}). \end{cases}$$

The HA Missing-Robustness Criterion

$$E\left(\left|\mathbf{X}^T\mathbf{D}^2\mathbf{X}\right|^{\frac{1}{p}}\right) = (1-\alpha)^n\left|\mathbf{X}^T\mathbf{X}\right|^{\frac{1}{p}} + \alpha(1-\alpha)^{n-1}\sum_{i=1}^{n}\left|\mathbf{X}_i^T\mathbf{X}_i\right|^{\frac{1}{p}}$$
$$+\alpha^2(1-\alpha)^{n-2}\sum_{\substack{i\neq j\\ i<j}}^{n}\left|\mathbf{X}_{ij}^T\mathbf{X}_{ij}\right|^{\frac{1}{p}}$$
$$+\alpha^3(1-\alpha)^{n-3}\sum_{\substack{i\neq j\neq k\\ i<j<k}}^{n}\left|\mathbf{X}_{ijk}^T\mathbf{X}_{ijk}\right|^{\frac{1}{p}}+\dots$$

Full Criterion Good and Bad

Pros

- Intuitive
- Naturally balances efficiency and robustness

Cons

- Computationally infeasible
- Based on D-, not I-criterion

Truncated HA Criterion

$$\mathrm{THA} = (1-\alpha)^n \left|\mathbf{X}^T\mathbf{X}\right|^{\frac{1}{p}} + \alpha(1-\alpha)^{n-1}\sum_{i=1}^{n}\left|\mathbf{X}_i^T\mathbf{X}_i\right|^{\frac{1}{p}}$$
$$+ \alpha^2(1-\alpha)^{n-2}\sum_{\substack{i\neq j\\ i<j}}^{n}\left|\mathbf{X}_{ij}^T\mathbf{X}_{ij}\right|^{\frac{1}{p}}$$
$$+ \alpha^3(1-\alpha)^{n-3}\sum_{\substack{i\neq j\neq k\\ i<j<k}}^{n}\left|\mathbf{X}_{ijk}^T\mathbf{X}_{ijk}\right|^{\frac{1}{p}}.$$

Truncated Criterion Good and Bad

Pros

- Computationally feasible
- Captures spirit of full criterion in small n, small α settings

Cons

- Still a D-based criterion
- Still significantly more computationally demanding than the D-criterion
- For larger n and/or α, may not provide robustness if larger number of runs are missing

Algorithm

Straightforward adaptation of coordinate exchange, which uses computational shortcuts.

$$\begin{aligned}\text{THA} = \left|\mathbf{X}^T\mathbf{X}\right|^{1/p} \Bigg(& (1-\alpha)^n + \alpha(1-\alpha)^{n-1}\sum_{i=0}^{n} R_i^{1/p} \\ & + \alpha^2(1-\alpha)^{n-2}\sum_{\substack{i\neq j\\ i<j}}^{n} R_{ij}^{1/p} \\ & + \alpha^3(1-\alpha)^{n-3}\sum_{\substack{i\neq j\neq k\\ i<j<k}}^{n} R_{ijk}^{1/p}\Bigg).\end{aligned}$$

Outline

Two Types of Models

Standard regression model: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$.

Models we consider:

$$f^T(\mathbf{x})\boldsymbol{\beta} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i$$

$$f^T(\mathbf{x})\boldsymbol{\beta} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 .$$

Three Types of Designs

1. Classical (fractional factorials; CCDs)
2. Optimal (D-optimal; I-optimal)
3. Missing-Robust (truncated HA criterion)

Measuring Design Quality for First-Order Models

Use the determinant of the information matrix, $\left|\mathbf{X}^T\mathbf{X}\right|$.

What if m observations are missing?

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

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:

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

If m observations are missing?

$$I_F(i,m) = \frac{I_n^*}{I_{n-m,i}},\ i = 1, 2, \ldots, \binom{n}{m},$$

How to Assess Impact of Missing Runs

Bottom line: Examine how much information 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

- classical, optimal, and missing-robust designs
- first- and second-order models
- various design sizes
- a small number of missing runs ($m = 0, 1, 2, 3$).

Outline

1. Motivation
2. Constructing Missing-Robust Designs
3. Comparing Designs in terms of Missing Observations
4. Results
5. Conclusions

First-order Models: $k = 5$

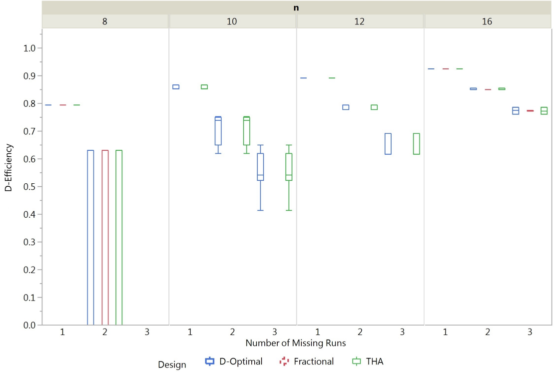
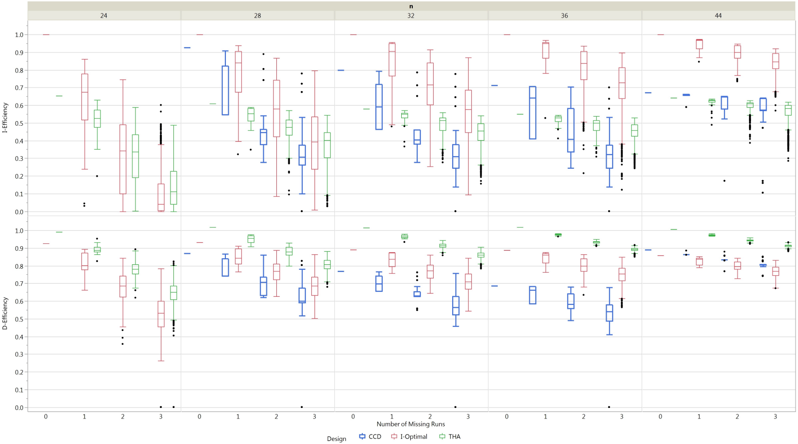


Figure: D-efficiencies for possible main effects designs, for $k = 5$ and $n = \{8, 10, 12, 16\}$, according to the number of missing runs.

Second-order Models: $k = 5$



Outline

Bottom Line

1. Optimal designs and classical designs have similar robustness properties, in terms of missing runs.

Bottom Line

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 (eg. 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)

Satirical headline: Professor celebrates landmark publication that will be carefully read by two people (h/t Maria Weese)

Reference

Smucker, B.J., Jensen, W., Wu, Z., and Wang, B. (2017). Robustness of Classical and Optimal Designs to Missing Observations. *Computational Statistics & Data Analysis*.
http://www.users.miamioh.edu/smuckebj/

Quick Plug

Quality Engineering

Submit your problems within the industrial statistics community

Print Email Tweet Share

Are you a practitioner of industrial statistics?

Can you think of a problem within the industrial statistics community?

As part of our new dedicated column within the *Quality Engineering* journal we will be looking at the issues you face within industrial statistics. Solutions will then be published as regular articles with the wider aim of improving collaboration between the industry and academia.

Would you like the industry to help solve your problems in industrial statistics? Then send us your problems, and they may be included in our dedicated open access column!

Motivation:

Industrial statistics research motivated by real-world problems and based on real data is extremely valuable, but relatively rare.

To encourage more and stronger collaboration between academia and industry, "Open Challenges in Industrial Statistics" will:

Column Editors

Column Editor:
Willis Jensen
W.L. Gore & Associates
wjensen@wlgore.com

Column Co-editor:
Byran Smucker
Miami University (OH)
smuckerb@miamioh.edu

Search for "open challenges quality engineering" or contact me (smuckerb@miamioh.edu).

Questions?

Extra Slide. Second-order Models: $k = 3$