Sizing Mixture (RSM) Designs for Adequate Precision via Fraction of Design Space (FDS)

Pat Whitcomb
Stat-Ease, Inc.
612.746.2036
fax 612.746.2056
pat@statease.com

Gary W. Oehlert
School of Statistics
University of Minnesota
612.625.1557
gary@stat.umn.edu

Sizing Mixture (RSM) Designs

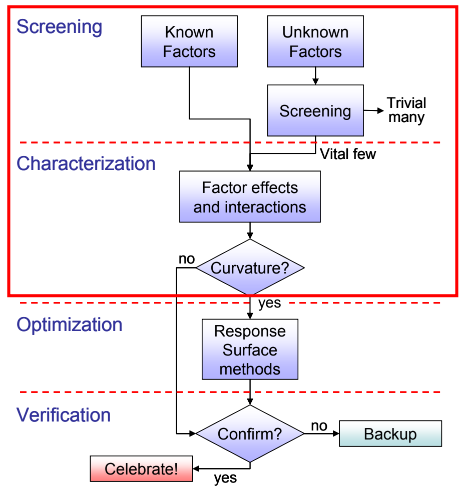
- **Review – Power to size factorial designs**
- Precision in place of power
  - Introduce FDS
- Sizing designs for precision
  - Two component mixture
  - Three component constrained mixture
- Sizing designs to detect a difference
- Summary

Sizing Factorial Designs

During screening and characterization the emphasis is on identifying factor effects.

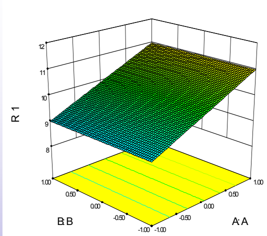
What are the important design factors?

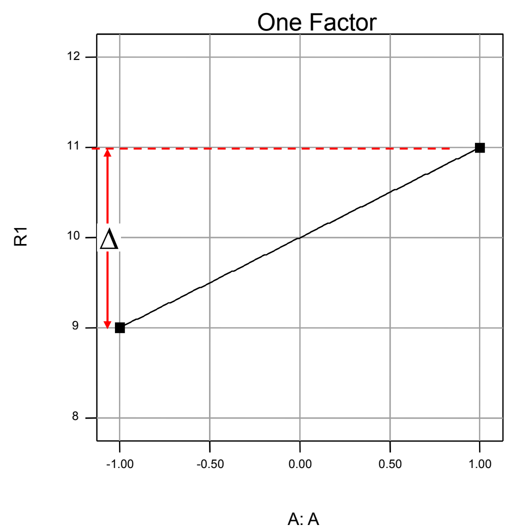
For this purpose power is an ideal metric to evaluate design suitability.



Factorial Design – Power

2^3 Full Factorial Δ=2 and σ=1





Factorial Design – Power

2^3 Full Factorial Δ=2 and σ=1

Leave Sigma and Delta fields blank to skip power calculation.

Responses: 1 (1 to 999) Edit model for power...

	Name	Units	Diff. to detect Delta("Signal")	Est. Std. Dev. Sigma("Noise")	Delta/Sigma (Signal/Noise Ratio)
	R1		2	1	2

Power is reported at a 5.0% alpha level to detect the specified signal/noise ratio.
Recommended power is at least 80%.

R1

Signal (delta) = 2.00 Noise (sigma) = 1.00 Signal/Noise (delta/sigma) = 2.00

A	B	C
57.2 %	57.2 %	57.2 %

The following three slides dig into how power is computed.

One Replicate of 2^3 Full Factorial

$C = (X^TX)^{-1}$ matrix

The design determines the standard error of the coefficient:

$$C = \begin{pmatrix} 0.125 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.125 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.125 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.125 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.125 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.125 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.125 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.125 \end{pmatrix}$$

$$\text{t-value}_i = \frac{\beta_i}{SE(\beta_i)} = \frac{\beta_i}{\sqrt{c_{ii}\hat{\sigma}^2}} = \frac{\beta_i}{\sqrt{(0.125)\hat{\sigma}^2}}$$

NonCentrality Parameter

2^3 Full Factorial Δ=2 and σ=1

The reference t distribution assumes the null hypothesis of Δ = 0. The noncentrality parameter (2.828) defines the t distribution under the alternate hypothesis of Δ = 2.

$$\text{noncentrality}_i = \frac{\beta_i}{\sqrt{c_{ii}\hat{\sigma}^2}} = \frac{\Delta_i/2}{\sqrt{c_{ii}\hat{\sigma}^2}}$$

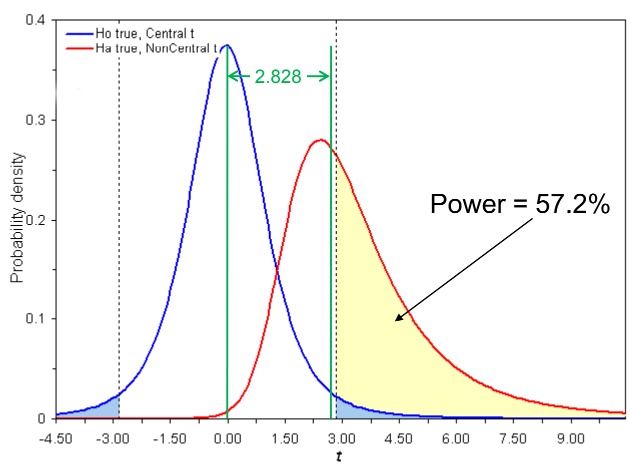
$$= \frac{1}{\sqrt{(0.125)(1)^2}}$$

$$= \frac{1}{0.3536} = 2.828$$

Factorial Design – Power

2^3 Full Factorial Δ=2 and σ=1

noncentral $t_{\alpha=0.05,df=4}$ with noncentrality parameter of 2.828



Factorial Design – Power

Two Replicates of 2^3 Full Factorial $\Delta=2$ and $\sigma=1$

Leave Sigma and Delta fields blank to skip power calculation.

Responses: 1 (1 to 999) Edit model for power...

Name	Units	Diff. to detect Delta("Signal")	Est. Std. Dev. Sigma("Noise")	Delta/Sigma (Signal/Noise Ratio)
R1		2	1	2

Power is reported at a 5.0% alpha level to detect the specified signal/noise ratio.
Recommended power is at least 80%.

R1

Signal (delta) = 2.00 Noise (sigma) = 1.00 Signal/Noise (delta/sigma) = 2.00

A	B	C
95.6 %	**95.6 %**	**95.6 %**

Two Replicates of 2^3 Full Factorial

$C = (X^TX)^{-1}$ matrix

The design determines the standard error of the coefficient:

$$C = \begin{pmatrix} 0.0625 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0625 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0625 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0625 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0625 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0625 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.0625 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.0625 \end{pmatrix}$$

$$\text{t-value}_i = \frac{\beta_i}{SE(\beta_i)} = \frac{\beta_i}{\sqrt{c_{ii}\hat{\sigma}^2}} = \frac{\beta_i}{\sqrt{(0.0625)\hat{\sigma}^2}}$$

NonCentrality Parameter

Two Replicates of 2^3 Full Factorial Δ=2 and σ=1

$$\text{noncentrality}_i = \frac{\beta_i}{\sqrt{c_{ii}\hat{\sigma}^2}} = \frac{\Delta_i/2}{\sqrt{c_{ii}\hat{\sigma}^2}}$$

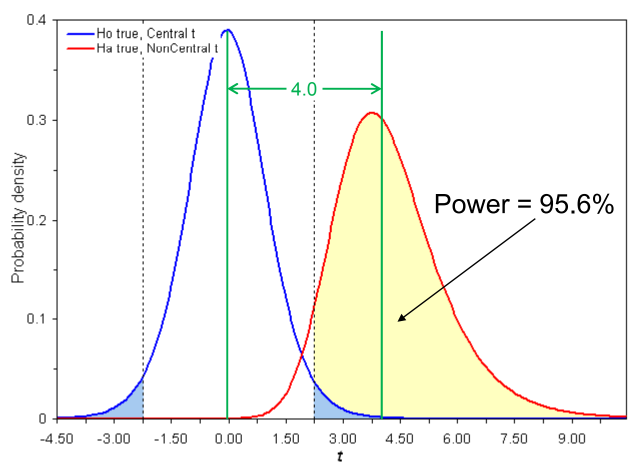
$$= \frac{1}{\sqrt{(0.0625)(1)^2}}$$

$$= \frac{1}{0.25} = 4.0$$

Factorial Design – Power

Two Replicates of 2^3 Full Factorial Δ=2 and σ=1

noncentral $t_{\alpha=0.05,df=12}$ with noncentrality parameter of 4.0



Conclusions

Factorial DOE
During screening and characterization (factorials) emphasis is on identifying factor effects. What are the important design factors? For this purpose power is an ideal metric to evaluate design suitability.

Sizing Mixture (RSM) Designs

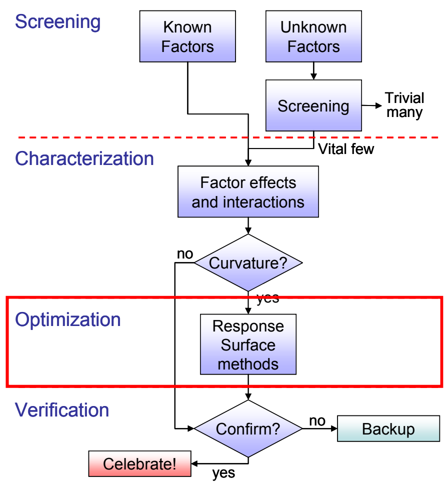
- Review – Power to size factorial designs
- **Precision in place of power**
  - **Introduce FDS**
- Sizing designs for precision
  - Two component mixture
  - Three component constrained mixture
- Sizing designs to detect a difference
- Summary

Sizing Response Surface Designs

When the goal is optimization emphasis is the fitted surface.

How well does the surface represent true behavior?

For this purpose precision (FDS) is an good metric to evaluate design suitability. *(Assuming model adequacy; i.e. insignificant lack of fit.)*



Design Properties

1. Estimate the designed for polynomial well.
2. Give sufficient information to allow a test for lack of fit.
   - ☑ Have more unique design points than coefficients in the model.
   - ☑ Provide an estimate of "pure" error.
3. Remain insensitive to outliers, influential values and bias from model misspecification.
4. Be robust to errors in control of the component levels.
5. Provide a check on model assumptions, e.g., normality of errors.
6. **Generate useful information throughout the region of interest, i.e., provide a good distribution of** $\sqrt{\mathrm{Var}(\hat{Y})/\sigma^2}$
7. Do not contain an excessively large number of trials.

FDS
Fraction of Design Space

Fraction of **D**esign **S**pace:

- Calculates the volume of the design space having a prediction variance (PV) less than or equal to a specified value.
- The ratio of this volume to the total volume of the design volume is the fraction of design space.
- Produces a single plot showing the cumulative fraction of the design space on the x-axis (from zero to one) versus the PV on the y-axis.

FDS – PV
Fraction of Design Space

Prediction Variance:

$$PV(x_0) = \frac{\text{var}(\hat{y}_0)}{s^2} = x_0^T (X^T X)^{-1} x_0$$

PV is a function of:

x_0 – the location in the design space (i.e. the x coordinates for all model terms).

X – the experimental design (i.e. where the runs are in the design space).

FDS – StdErr
Fraction of Design Space

Prediction standard error of the expected value:

$$PV\left(x_0\right) = \frac{\operatorname{var}\left(\hat{\bar{y}}_0\right)}{s^2} = x_0^T\left(X^T X\right)^{-1} x_0$$

$$StdErr\left(x_0\right) = \frac{s_{\hat{\bar{y}}_0}}{s} = \sqrt{PV\left(x_0\right)}$$

FDS Plot
Fraction of **D**esign **S**pace

1. Pick random points in the design space.
2. Calculate the standard error of the expected value

$$\frac{SE_{\hat{\bar{y}}_0}}{s} = \sqrt{x_0^T\left(X^T X\right)^{-1} x_0}$$

3. Plot the standard error as a fraction of the design space.