Design and Analysis of Multi-Factored Experiments

Response Surface Methodology (RSM)

Introduction to Response Surface Methodology (RSM)

• Best and most comprehensive reference:

• Best software:
  – Design-Expert Version 8, Statease Inc.
  – Available at www.statease.com
  – Minitab also has DOE and RSM capabilities
RSM: Introduction

• Primary focus of previous discussions is factor screening
  – Two-level factorials, fractional factorials are widely used
• RSM dates from the 1950s (Box and Wilson, 1951)
• Early applications in the chemical industry
• Currently RSM is widely used in quality improvement, product design, uncertainty analysis, etc.

Objective of RSM

• RSM is a collection of mathematical and statistical techniques that are useful for modeling and analysis in applications where a response of interest is influenced by several variables and the objective is to optimize the response.
  • Optimize \(\rightarrow\) maximize, minimize, or getting to a target.
  • Or, where a nonlinear model is warranted when there is significant curvature in the response surface.
Uses of RSM

• To determine the factor levels that will simultaneously satisfy a set of desired specification (e.g. model calibration)
• To determine the optimum combination of factors that yield a desired response and describes the response near the optimum
• To determine how a specific response is affected by changes in the level of the factors over the specified levels of interest

Uses of RSM (cont)

• To achieve a quantitative understanding of the system behavior over the region tested
• To find conditions for process stability = insensitive spot (robust condition)
• To replace a more complex model with a much simpler second-order regression model for use within a limited range \( \rightarrow \) replacement models, meta models, or surrogate models. E.g. Replacing a FEM with a simple regression model.
Example

Suppose that an engineer wishes to find the levels of temperature \(x_1\) and feed concentration \(x_2\) that maximize the yield \(y\) of a process. The yield is a function of the levels of \(x_1\) and \(x_2\), by an equation:

\[
Y = f(x_1, x_2) + \varepsilon
\]

If we denote the expected response by

\[
E(Y) = f(x_1, x_2) = \eta
\]

then the surface represented by:

\[
\eta = f(x_1, x_2)
\]

is called a **response surface**.

The response surface maybe represented graphically using a contour plot and/or a 3-D plot. In the contour plot, lines of constant response \(y\) are drawn in the \(x_1, x_2\) plane.
These plots are of course possible only when we have two factors.

With more than two factors, the optimal yield has to be obtained using numerical optimization methods.

In most RSM problems, the form of the relationship between the response and the independent variables is unknown. Thus, the first step in RSM is to find a suitable approximation for the true relationship between Y and the X’s.
If the response is well modeled by a linear function of the independent variables, then the approximating function is the first-order model (linear):

\[ Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \]

This model can be obtained from a $2^k$ or $2^{k-p}$ design.

If there is curvature in the system, then a polynomial of higher degree must be used, such as the second-order model:

\[ Y = \beta_0 + \Sigma \beta_i x_i + \Sigma \beta_{ii} x_i^2 + \Sigma \Sigma \beta_{ij} x_i x_j + \varepsilon \]

This model has linear + interaction + quadratic terms.

• Many RSM problems utilize one or both of these approximating polynomials. The response surface analysis is then done in terms of the fitted surface. The 2\textsuperscript{nd} order model is nearly always adequate if the surface is “smooth”.

• If the fitted surface is an adequate approximation (high $R^2$) of the true response function, then analysis of the fitted surface will be approximately equivalent to analysis of the actual system (within bounds).
Types of functions

- Figures 1a through 1c on the following pages illustrate possible behaviors of responses as functions of factor settings. In each case, assume the value of the response increases from the bottom of the figure to the top and that the factor settings increase from left to right.
• If a response behaves as in Figure 1a, the design matrix to quantify that behavior need only contain factors with two levels -- low and high.
• This model is a basic assumption of simple two-level factorial and fractional factorial designs.
• If a response behaves as in Figure 1b, the minimum number of levels required for a factor to quantify that behavior is three.

• One might logically assume that adding center points to a two-level design would satisfy that requirement, but the arrangement of the treatments in such a matrix confounds all quadratic effects with each other.

• **While a two-level design with center points cannot estimate individual pure quadratic effects, it can detect them effectively.**

• A solution to creating a design matrix that permits the estimation of simple curvature as shown in Figure 1b would be to use a three-level factorial design. Table 1 explores that possibility.

• Finally, in more complex cases such as illustrated in Figure 1c, the design matrix must contain at least four levels of each factor to characterize the behavior of the response adequately.
Table 1: 3 level factorial designs

<table>
<thead>
<tr>
<th>No. of factors</th>
<th># of combinations(3^k)</th>
<th>Number of coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>81</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>243</td>
<td>21</td>
</tr>
<tr>
<td>6</td>
<td>729</td>
<td>28</td>
</tr>
</tbody>
</table>

- The number of runs required for a $3^k$ factorial becomes unacceptable even more quickly than for $2^k$ designs.
- The last column in Table 1 shows the number of terms present in a quadratic model for each case.

Problems with 3 level factorial designs

- With only a modest number of factors, the number of runs is very large, even an order of magnitude greater than the number of parameters to be estimated when $k$ isn't small.
- For example, the absolute minimum number of runs required to estimate all the terms present in a four-factor quadratic model is 15: the intercept term, 4 main effects, 6 two-factor interactions, and 4 quadratic terms.
- The corresponding $3^k$ design for $k = 4$ requires 81 runs.
• Considering a fractional factorial at three levels is a logical step, given the success of fractional designs when applied to two-level designs.
• Unfortunately, the alias structure for the three-level fractional factorial designs is considerably more complex and harder to define than in the two-level case.
• Additionally, the three-level factorial designs suffer a major flaw in their lack of ‘rotatability’
• More on ‘rotatability’ later.

Sequential Nature of RSM
• Before going on to economical designs to fit second-order models, let’s look at how RSM is carried out in general.
• RSM is usually a sequential procedure. That is, it done in small steps to locate the optimum point, if that’s the objective. This is not always the only objective.
• The analogy of climbing a hill is appropriate here (especially if it is a very foggy day)!
Sequential Nature of RSM (continue)

- When we are far from the optimum (far from the peak) there is little curvature in the system (slight slope only), then first-order model will be appropriate.
- The objective is to lead the experimenter rapidly and efficiently to the general vicinity of the optimum.
- Once the region of the optimum has been found, a more elaborate model such a second-order model may be employed, and an analysis performed to locate the optimum.
• The eventual objective of RSM is to determine the optimum operating conditions for the system or to determine a region of the factor space in which operating specifications are satisfied.

• The word “Optimum” in RSM is used in a special sense. The “hill climbing” procedures of RSM guarantee convergence to a local optimum only.

• In terms of experimental designs, when we are far from optimum, a simple $2^k$ factorial experiment would allow us to fit a first-order model. As we get nearer to the peak, we can check for curvature by adding center-points to the $2^k$ factorial.

• If curvature is significant, we may now be in the vicinity of the peak and we use a more elaborate design (e.g. a CCD) to fit a second-order model to “capture” the optimum.
Method of Steepest Ascent

- The **method of steepest ascent** is a procedure for moving sequentially along the path of steepest ascent (PSA), that is, in the direction of the maximum increase in the response. If minimization is desired, then we are talking about the **method of steepest descent**.

- For a first-order model, the contours of the response surface is a series of parallel lines. The direction of steepest ascent is the direction in which the response \( y \) increases most rapidly. This direction is normal (perpendicular) to the fitted response surface contours.
Path of Steepest Ascent (PSA)

- The PSA is usually the line through the center of the region of interest and normal to the fitted surface contours.
- The steps along the path are proportional to the regression coefficients \( \{\beta_i\} \). The actual step size would depend on the experimenter’s knowledge of the process or other practical considerations.

For example, consider the first-order model:

\[
y = 40.00 + 0.775 x_1 + 0.325 x_2
\]

For steepest ascent, we move 0.775 unit in the \( x_1 \) direction for every 0.325 unit in the \( x_2 \) direction.

Thus the PSA passes through the center (0, 0) and has a slope of 0.375/0.775.
If say 1 unit of $x_1$ is actually equal to 5 minutes in actual units, and 1 unit of $x_2$ is actually equal to 5 °F,

the PSA are $\Delta x_1 = 1.00$ and

$\Delta x_2 = (0.375/0.775)$

$\Delta x_2 = 0.42 = 2.1$ °F.

Therefore, you will move along the PSA by increasing time by 5 minutes and temperature by 2 °F. An actual observation on yield will be determined at each point.

- Experiments are then conducted along the PSA until no further increase in the response is observed.
- Then a new first-order model may be fit, a new direction of steepest ascent determined, and further experiments conducted in that direction until the experimenter feels that the process is near the optimum (peak of hill is within grasp!).
- The steepest ascent would terminate after about 10 steps with an observed response of about 80%. Now we move on to the next step.
- Fit another first-order model with a new center (where step 10 is) and check whether there is a new PSA.
- Repeat until peak is near.
- See flowchart on the next slide.
Steps in RSM

- Fit linear model/planar models using two-level factorials
- From results, determine PSA (Descent)
- Move along path until no improvement occurs
- Repeat steps 1 and 2 until near optimal (change of direction is possible)
- Fit quadratic model near optimal in order to determine curvature and find peak. This phase is often called “method of local exploration”
- Run confirmatory tests
• With well-behaved functions with a single peak or valley, the above procedure works very well. It becomes more difficult to use RSM or any other optimization routine when the surface has many peaks, ridges, and valleys.
Multiple Objectives

• With more than 2 factors, it is more difficult to determine where the optimal is. There may be several possible “optimal” points and not all are desirable. Whatever the final choice of optimal factor levels, common sense and process knowledge must be your guide.

• It is also possible to have more than one response variable with different objectives (sometimes conflicting). For these cases, a weighting system may be used to for the various objectives.

Methods of Local Exploration

• The method of steepest ascent, in addition to fitting first-order model, must provide additional information that will eventually identify when the first-order model is no longer valid.

• This information can come only from additional degrees of freedom which are used to measure “lack of fit” in some way.

• This means additional levels and extra data points.

• It is rare to go more than 5 levels for even the most complex response surfaces.
Consider the 2nd order model:

\[
Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \\
+ \beta_{11} x_1^2 + \beta_{22} x_2^2 + \ldots + \beta_{kk} x_k^2 \\
+ \beta_{12} x_1 x_2 + \ldots + \beta_{1k} x_1 x_k + \ldots + \beta_{23} x_2 x_3 \\
+ \ldots + \beta_{k-1,k} x_{k-1} x_k + \epsilon \quad \text{------- EQN (1)}
\]

To be able to fit a 2nd order model like EQN (1), there must be at least three levels and enough data points.

Designs for fitting 2nd order models

- Two very useful and popular experimental designs that allow a 2nd order model to be fit are the:
  - Central Composite Design (CCD)
  - Box-Becken Design (BBD)
- Both designs are built up from simple factorial or fractional factorial designs.
Central Composite Design (CCD)

- Each factor varies over five levels
- Typically smaller than Box-Behnken designs
- Built upon two-level factorials or fractional factorials of Resolution V or greater
- Can be done in stages → factorial + centerpoints + axial points
- Rotatable
General Structure of CCD

- $2^k$ Factorial + 2k Star or axial points + $n_c$ Centerpoints

- The factorial part can be a fractional factorial as long as it is of Resolution V or greater so that the 2 factor interaction terms are not aliased with other 2 factor interaction terms.

- The “star” or “axial” points in conjunction with the factorial and centerpoints allows the quadratic terms ($\beta_{ii}$) to be estimated.

Generation of a CCD

Factorial points + centerpoints

Axial points

L. M. Lye  DOE Course  43

L. M. Lye  DOE Course  44
Axial points are points on the coordinate axes at distances “\( \alpha \)” from the design center; that is, with coordinates: For 3 factors, we have \( 2k = 6 \) axial points like so:

\[
(+\alpha, 0, 0), (-\alpha, 0, 0), (0, +\alpha, 0), (0, -\alpha, 0), (0, 0, +\alpha), (0, 0, -\alpha)
\]

The “\( \alpha \)” value is usually chosen so that the CCD is rotatable.

At least one point must be at the design center \((0, 0, 0)\). Usually more than one to get an estimate of “pure error”. See earlier 3-D figure.

If the “\( \alpha \)” value is 1.0, then we have a face-centered CCD \( \rightarrow \) Not rotatable but easier to work with.
Values of $\alpha$ for CCD to be rotatable

<table>
<thead>
<tr>
<th>k=2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.414</td>
<td>1.682</td>
<td>2.000</td>
<td>2.378</td>
<td>2.828</td>
<td>3.364</td>
</tr>
</tbody>
</table>

The $\alpha$ value is calculated as the $4^{th}$ root of $2^k$.

For a rotatable design the variance of the predicted response is constant at all points that are equidistant from the center of the design.

Types of CCDs

The diagrams illustrate the three types of central composite designs for two factors. Note that the CCC explores the largest process space and the CCI explores the smallest process space. Both the CCC and CCI are rotatable designs, but the CCF is not. In the CCC design, the design points describe a circle circumscribed about the factorial square. For three factors, the CCC design points describe a sphere around the factorial cube.
Box-Behnken Designs (BBD)

- The Box-Behnken design is an independent quadratic design in that it does not contain an embedded factorial or fractional factorial design.
- In this design the treatment combinations are at the midpoints of edges of the process space and at the center.
- These designs are rotatable (or near rotatable) and require 3 levels of each factor.
- The designs have limited capability for orthogonal blocking compared to the central composite designs.

BBD - summary

- Each factor is varied over three levels (within low and high value)
- Alternative to central composite designs which requires 5 levels
- BBD not always rotatable
- Combinations of 2-level factorial designs form the BBD.
A 3-Factor BBD with 1 centerpoint

<table>
<thead>
<tr>
<th>Runs</th>
<th>x_1</th>
<th>x_2</th>
<th>x_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>6</td>
<td>-1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Brief Comparison of CCD and BBD

With one centerpoint, for
k = 3, CCD requires 15 runs; BBD requires 13 runs
k = 4, CCD requires 25 runs; BBD also requires 25 runs
k = 5, CCD requires 43 runs; BBD requires 41 runs

but, for CCD we can run a 2^{5-1} FFD with Resolution V. Hence we need only 27 runs.

In general CCD is preferred over BBD. See separate handout comparing CCD and BBD in more detail.
Analysis of the fitted response surface

• The fitted response surface can take on many shapes.
• For 2 or less dimensions, we can plot the response against the factor(s) and graphically determine where the optimal response is.
• We can also tell from the contour plots or 3-D plots whether we have a maximum, minimum, or a saddle point. These points are stationary points.

Types of stationary points

a) Maximum point; b) Minimum point; c) Saddle point
With more than 2 factors

- For more than 2 factors, we need to use numerical methods to tell what kind of stationary point we have.
- In some cases, even this fails.
- The levels of the k factors at which the response is optimal can be determined for the unconstrained case by simple calculus.

When k=1

Consider the 2nd order prediction model with k =1:

\[ \hat{y} = b_0 + b_1x + b_2x^2 \]

Provided that \( b_2 \) is not zero, the optimum response is obtained by:

\[ \frac{dy}{dx} = b_1 + 2b_2x = 0 \]

Giving:

\[ x = -\frac{b_1}{2b_2} \]
For $k > 1$

In the case of $k > 1$, we can write the 2nd order equation

\[
y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k + b_{11} x_1^2 + b_{22} x_2^2 + \ldots + b_{kk} x_k^2 + b_{12} x_1 x_2 + \ldots + b_{1k} x_1 x_k + \ldots + b_{k-1,k} x_{k-1} x_k
\]

in a more convenient matrix form as:

\[
y = b_0 + x' b + x' B x
\]

Where:

\[
\begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_k
\end{bmatrix}
\]

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_k
\end{bmatrix}
\]

\[
B = \begin{bmatrix}
  b_{11} & \frac{1}{2} b_{12} & \cdots & \frac{1}{2} b_{1k} \\
  \frac{1}{2} b_{12} & b_{22} & \cdots & \frac{1}{2} b_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  \frac{1}{2} b_{1k} & \frac{1}{2} b_{2k} & \cdots & b_{kk}
\end{bmatrix}
\]
Provided that matrix $B$ is not singular, the 2nd order model has a stationary point (i.e., a point at which first partial derivatives with respect to $x_1$, $x_2$, ..., $x_k$ are all 0) given by:

$$x_0 = -\frac{1}{2}B^{-1}b$$

Depending on the nature of $B$, the stationary point will be either a minimum, a maximum, or a saddle point of the fitted surface. Moving away from a saddle point in some directions produces an increase in the response, while moving away in other directions produces a decrease in the response.

Characterization of the stationary point

If $B$ is positive definite $\Rightarrow$ all eigenvalues are positive $\Rightarrow$ minimum pt.

If $B$ is negative definite $\Rightarrow$ all eigenvalues are negative $\Rightarrow$ maximum pt.

If $B$ is indefinite $\Rightarrow$ eigenvalues are positive and negative $\Rightarrow$ saddle pt.

Eigenvalues of matrix $B$ can be obtained using MATLAB
Be aware that the $x_0$ obtained are random variables and have associated uncertainty with them as are the eigenvalues and matrix $B$.

There will be situations when an unconstrained optimum will not be useful (when there is a saddle point). We need to consider a constraint that forces us to stay within the experimental region. The procedure that has been developed for this is called **ridge analysis** (see Myers and Montgomery). Need to use Lagrange multipliers here for the optimization.

There are also other methods for solving optimization problems constrained or unconstrained that do not require the taking of partial derivatives

→ Direct Optimization procedures. E.g. Nelder-Mead simplex search procedure, or by Monte Carlo simulation.

RSM are now done mainly by software except for simple cases.
Other Aspects of Response Surface Methodology

• Robust parameter design and process robustness studies
  – Find levels of controllable variables that optimize mean response and minimize variability in the response transmitted from “noise” variables
  – Original approaches due to Taguchi (1980s)
  – Modern approach based on RSM

• Experiments with mixtures
  – Special type of RSM problem
  – Design factors are components (ingredients) of a mixture
  – Response depends only on the proportions
  – Many applications in product formulation

Designs for computer experiments

• Much developments of sophisticated engineering designs, analysis, and products are now carried out by high-powered computer simulations.
• Some of these sophisticated programs require either expensive computing resources or computer time.
• Hence simplifying the model by means of a meta model or replacement model often makes more sense. Done properly using DOE methods also helps to understand the complex model a little better.
• If the objective is to estimate a polynomial transfer function, traditional RSMs such as CCD and BBD have been used with some success.

• However, when analyzing data from computer simulations, we must keep in mind that the true model will only be approximated by RSM.

• The RSM metamodel will not only fall short in the form of the model, but also in the number of factors.

• Therefore, predictions will only be good within the ranges of the factors specified and will exhibit systematic error, or bias.

• The systematic error is what will be measured in the residual – not the normal variations observed from a random physical process.

• Despite these circumstances, much of the standard statistical analyses remain relevant, including model-fit such as the Prediction $R^2$.

• However, the p-values will not be accurate estimates of risks associated with the overall model or any of its specific terms.

• The goal of fitting a RSM to deterministic computer simulated data is for a perfect fit so that there is no systematic error.
Check list for quality of fit of designs for RSM

- **Generate information throughout the region of interest.**
- Ensure the fitted value be as close as possible to the true value.
- Give good detectability of lack of fit.
- Allow designs of increasing order to be built up sequentially.
- Require a minimum number of runs.
- Choose unique design points in excess of the number of coefficients in the model.
- Remain insensitive to influential values and bias from model.
- Allows one to fit a variety of models.

---

Newer DOE for Computer Experiments

- Computer models of actual or theoretical physical systems can take many forms and different levels of granularity of representation of the physical system.
- Models are often very complicated and constructed with different levels of fidelity such as the detailed physics-based model as well as more abstract and higher level models with less detailed representation.
- A physics-based model may be represented by a set of equations including linear, nonlinear, ordinary, and partial differential equations.
• In view of the complex and nonlinear nature of modern computer models, the classical RSM approaches usually do not provide adequate coverage of the experimental area to provide an accurate metamodel.

• To find a high quality metamodel, choosing a good set of “training” data becomes an important issue for computer simulation.

• Efficient “Space-Filling” designs are able to generate a set of sample points that capture the maximum information between the input-output relationships.

• E.g. Uniform Designs and Latin hypercube sampling are two such designs.

• http://www.math.hkbu.edu.hk/UniformDesign/

Example of a 20 run, 3 factor, 4 level Uniform Design Uq⁴ (Centered L₂)

Levels must be equally spaced.

4 levels allow a cubic equation to be fitted.

Correlations: A, B, C

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Correlations:

A    B
-0.000 1.000

B    C
0.040 0.000
0.867 1.000

C    1
Number of parameters for various models

<table>
<thead>
<tr>
<th># FACTORS</th>
<th>LINEAR</th>
<th>QUADRATIC</th>
<th>CUBIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>15</td>
<td>35</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>21</td>
<td>56</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>28</td>
<td>84</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>36</td>
<td>120</td>
</tr>
</tbody>
</table>

N = # of parameters + 4 additional points.

- How to find the best suited metamodel is another key issue in computer experiments.
- Techniques include: kriging models, polynomial regression models, local polynomial regression, multivariate splines and wavelets, and neural networks have been proposed.
- Therefore, design and modelling are two key issues in computer experiments.
- Most of these techniques are outside of statistics although knowledge of classical DOE and RSM certainly helps in understanding these new techniques.

See papers by Kleijnen et al for more details.
12th Annual Golfing Challenge – 3 holes

- Conduct an experiment using the golfing toy and obtain a prediction equation for the toy for use in a 3-hole golf championship to be played using the toy in the Faculty Lounge on Saturday, November 17th.
- [Hint: Use a face-centered CCD RS design or BBD RS design]
- Team with the least total number of strokes over 3 holes wins 5 extra marks and bragging rights!
- Each team must also submit a report of your experimental design.