# Experimental Design

Peter J. Crickmore Ph.D., P.Eng.

CEI Inc.

# Topics Covered

- Introduction
- Full 2 Level Factorial Designs
- Fractional 2 Level Factorial Designs (Generators, Resolution and Plackett Burman Designs)
- Variances, Significance and Error Bounds
- Response Surface Designs (Full 3 Level, Box Behnken, and Central Composite Designs)
- Mixtures and Scheffé Designs
- Sequential Designs, Evolutionary Operation, and the Simplex Method

# What is Experimental Design?

- Designing experiments to yield the most information from the fewest runs
- Isn’t this obvious?  Shouldn’t most people with a science or engineering background know this stuff intuitively?

# History

- Agricultural field trial work of Fisher* in the 1920s
- Chemical industry work of George Box, starting in 1951 with the publication of his work term notes from ICI**

*R.A. Fisher, ”Studies in Crop Variation. II. The Manurial Response of Different Potato Varieties”, J. Agricultural Sci., 13, 311-320 (1923)

** G.E.P. Box and K.B. Wilson, ”On the experimental attainment of optimal conditions”, J. Roy. Statist. Soc., B13,1-45 (1951)

# Ronald A. Fisher

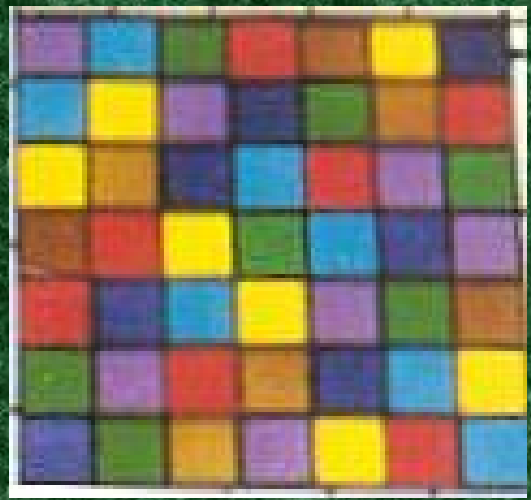
CEI Inc.

# George E.P. Box

CEI Inc.


Fisher's window is a 7x7 Latin square for yield studies on potatoes.

Memorial windows for Venn and Fisher, Caius College, Cambridge



# N-space and Real Space

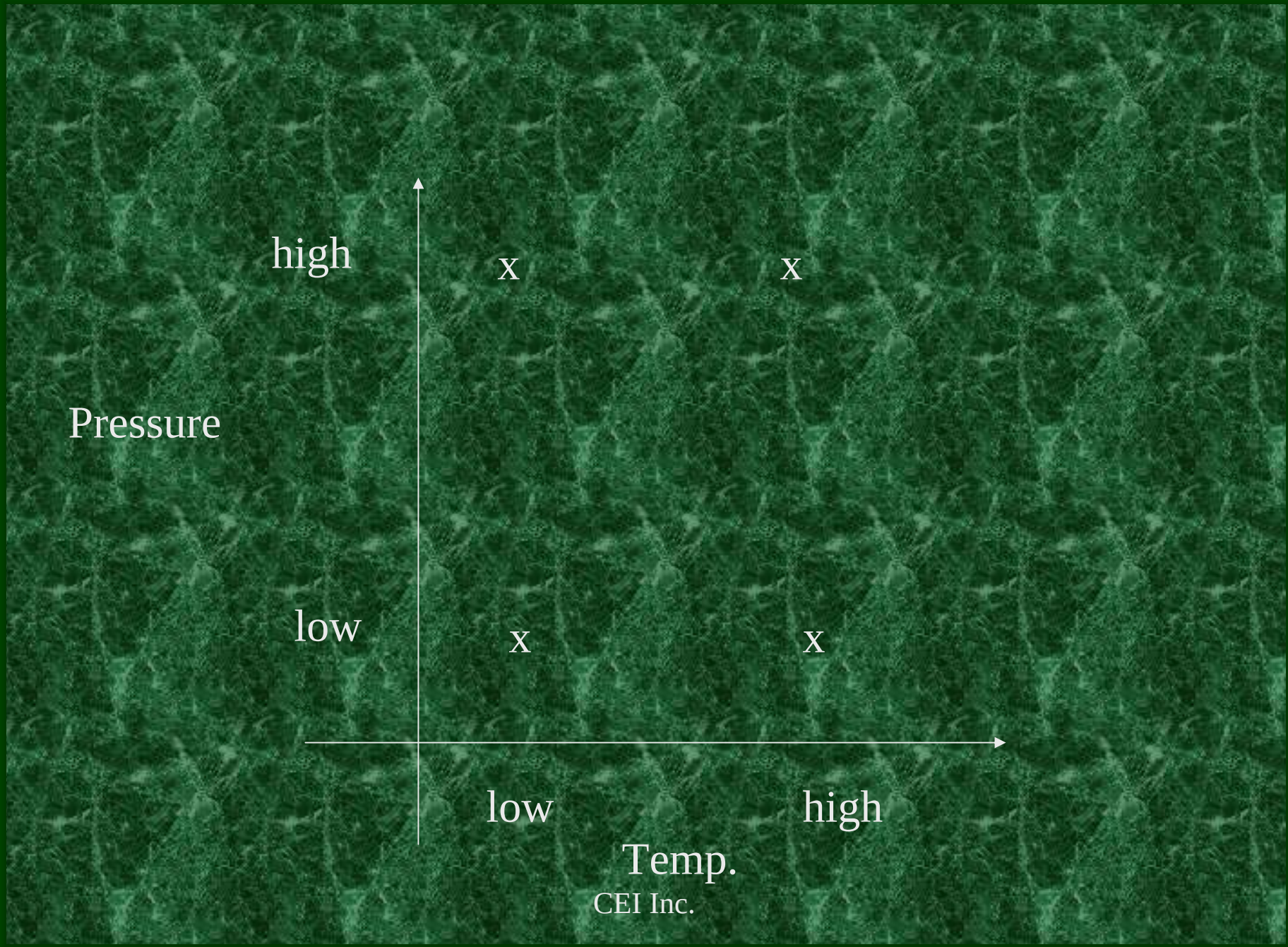
- It is often useful to move back and forth between sampling for a spatial function of x and y (or, of x, y and z) and testing a function of several other independent variables (temperature, pressure, concentration of reactants, etc.)

# Two Basic Types of Design

- One that identifies influential variables or the “effects” different variables have
- One that determines approximate functional form of a relationship or the “response surface”

# Two Level Factorial Designs

- If we are to determine which variables influence a result, usually use a two level factorial design
- For example, determine if temperature and pressure influence the specific volume of ammonia -- will look at P and T at “high” and “low” levels


high
x
x
Pressure
low
x
x
low
high
Temp.

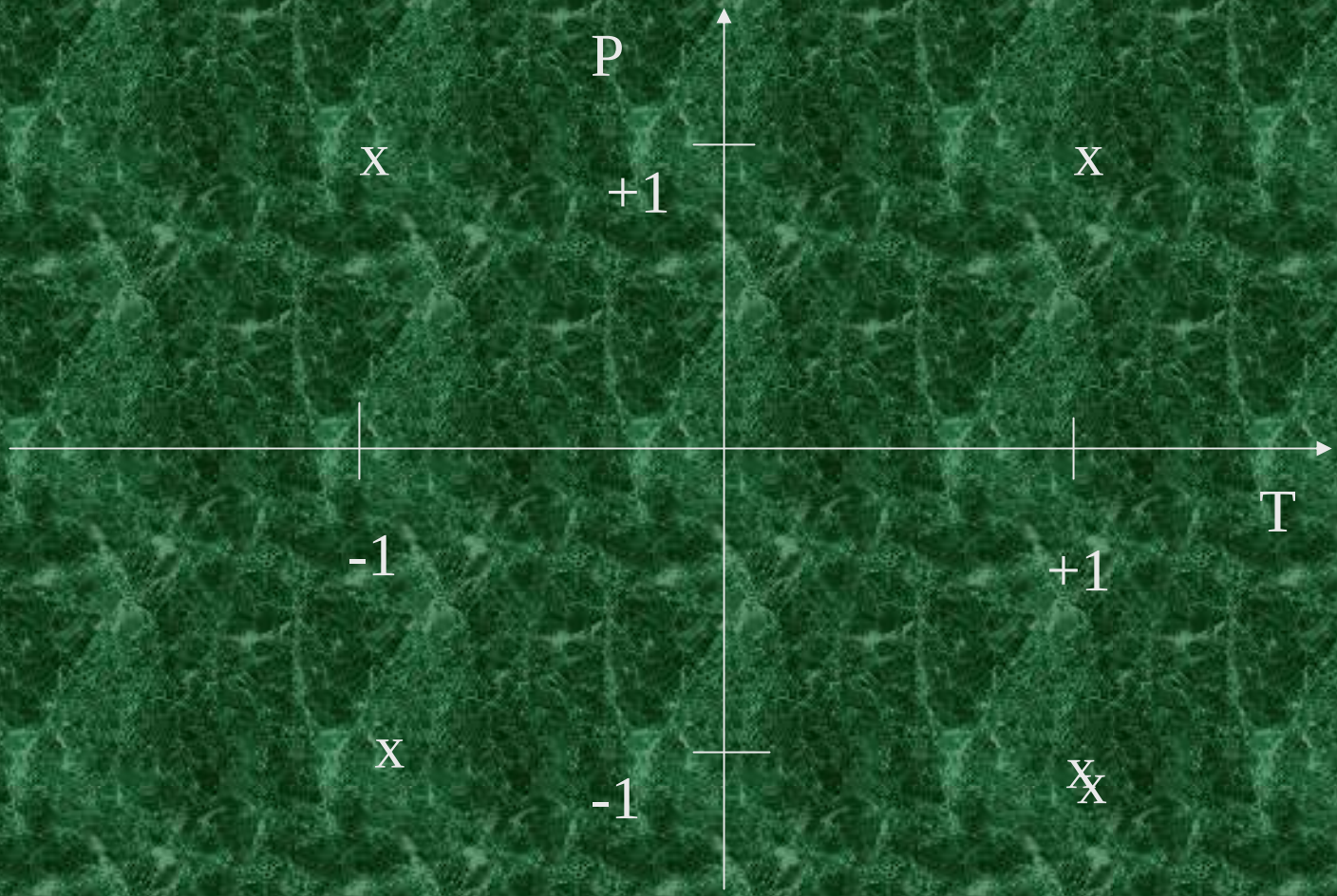
# Specific Volume of Ammonia

- Four runs are needed to get all combinations of “high” and “low”
- Usual to use coded independent variables

$$X_c = 2\,(x - x_{av})/(x_{high} - x_{low})$$

# Coding

- For example, if the temperature ranges from 40 $^{o}$C to 140 $^{o}$C
  - If T = 40 $^{o}$C, $T_c$ = -1
  - If T = 140 $^{o}$C, $T_c$ = +1

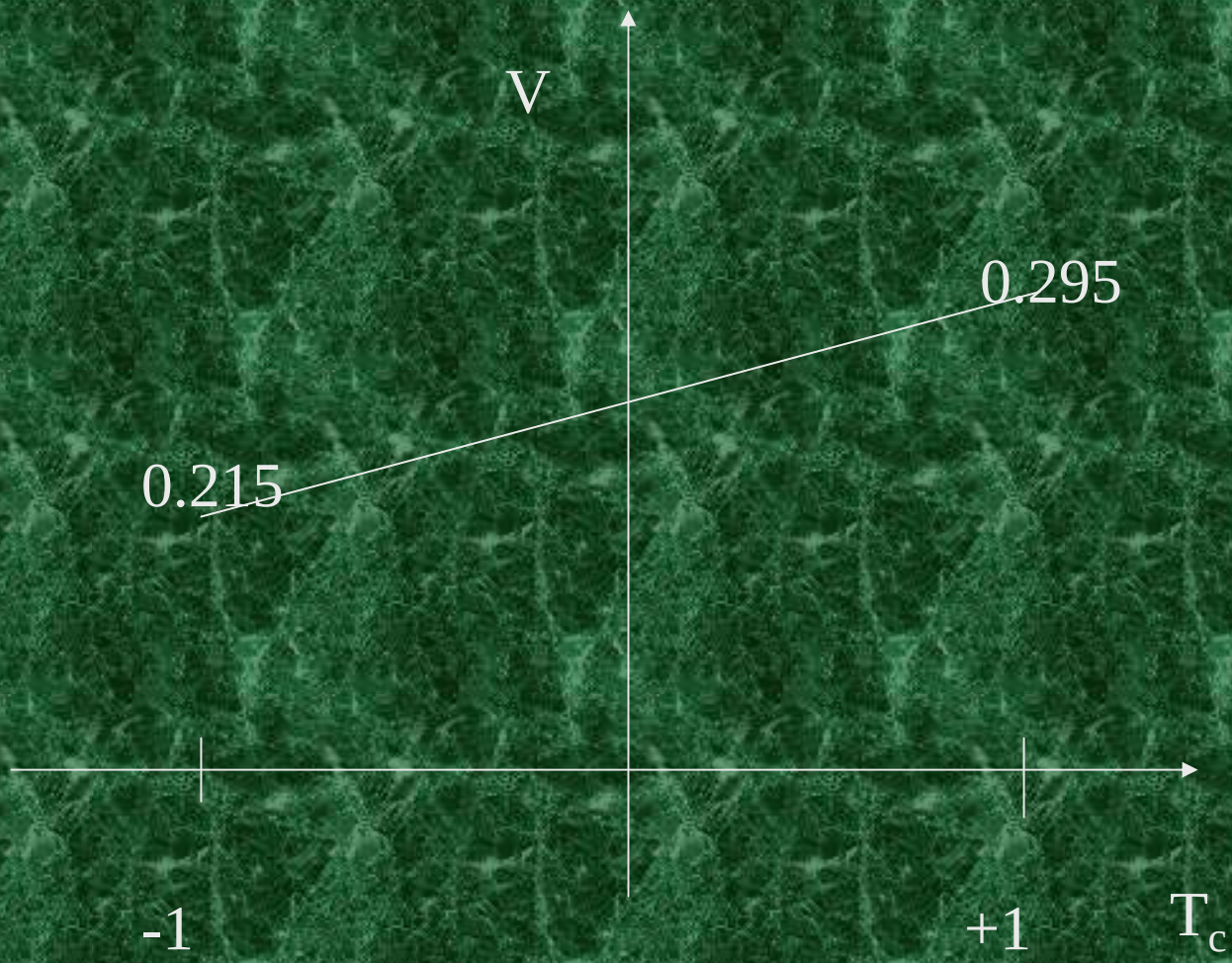



Full Two Level Factorial Design in Two Factors

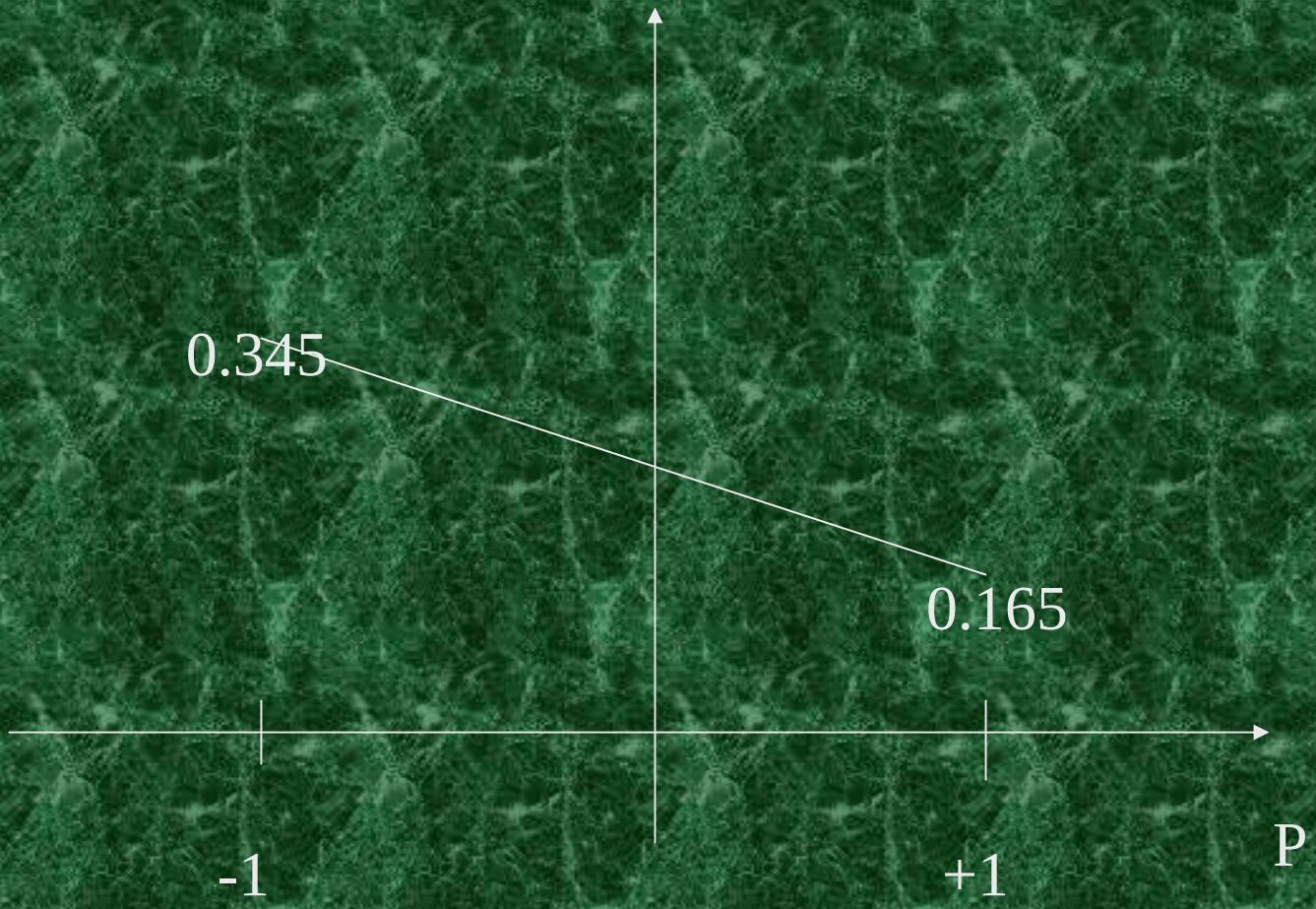
# Experimental Runs

| Order | T, $^{o}C$ | P, kPa | $T_{coded}$ | $P_{coded}$ | V, $m^3/kg$ |
|---|---|---|---|---|---|
| | | | | | |
| 3 | 40 | 500 | -1 | -1 | 0.29 |
| 2 | 40 | 1000 | -1 | +1 | 0.14 |
| 4 | 140 | 500 | +1 | -1 | 0.40 |
| 1 | 140 | 1000 | +1 | +1 | 0.19 |

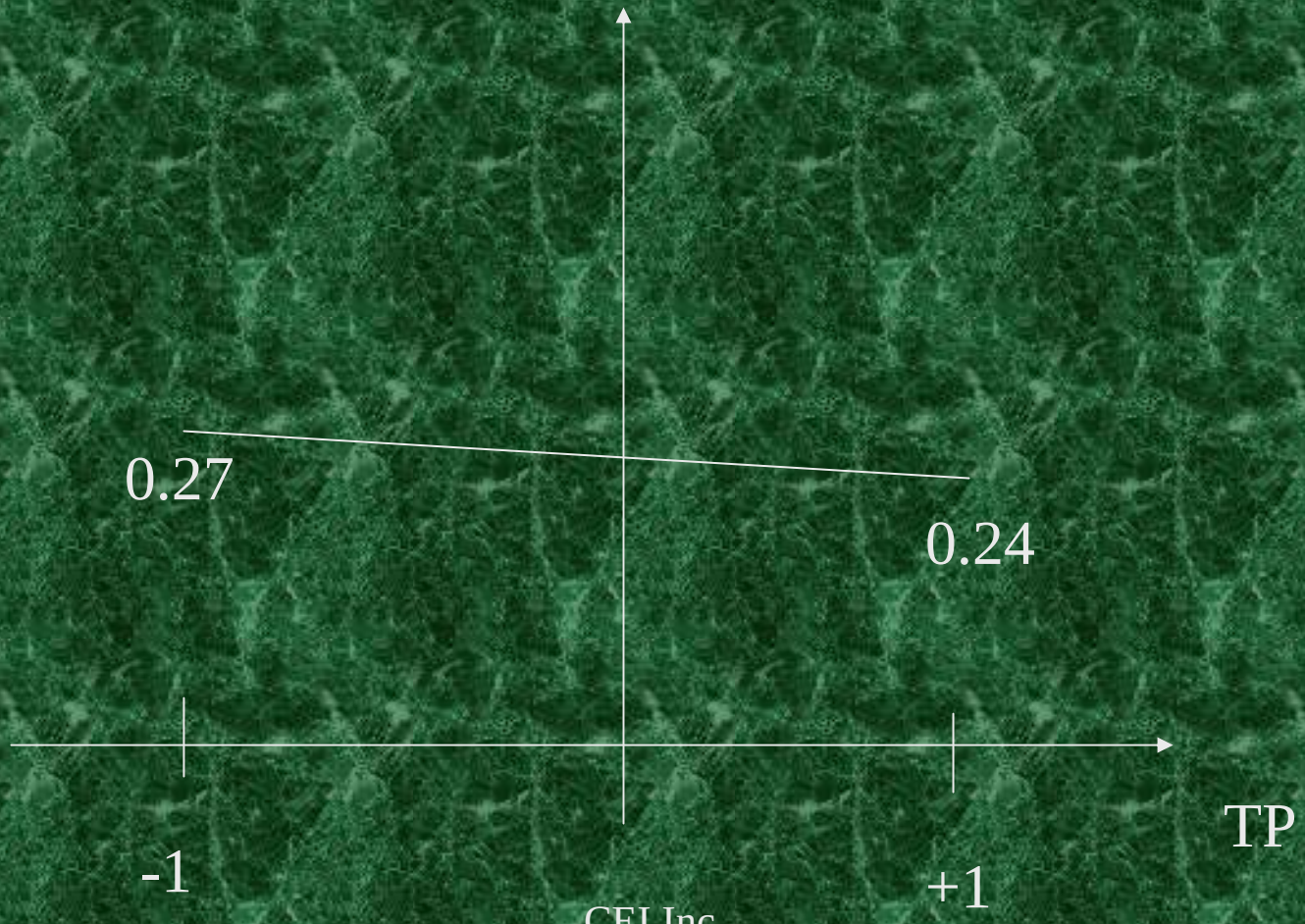
## Effect of Temperature?



# Effect of Pressure?



# Effect of T, P Interaction?



# Effect of $T^2$? $P^2$?

In 2 level factorial designs, the squared terms are always equal to 1.

# Empirical Models

- It is often preferable to fit to some sort of empirical model
- It can be shown that the most general polynomial model that can be estimated from a full 2 level factorial design in two factors is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Our data can be written out as

$$\beta_0 - \beta_1 - \beta_2 + \beta_{12} = 0.29$$

$$\beta_0 - \beta_1 + \beta_2 - \beta_{12} = 0.14$$

$$\beta_0 + \beta_1 - \beta_2 - \beta_{12} = 0.40$$

$$\beta_0 + \beta_1 + \beta_2 + \beta_{12} = 0.19$$

i.e. four equations in four unknowns, so there should be a solution.

The solution is easier if we use matrices (really).

$$\underline{y} = \underline{X}\,\underline{\beta}$$

$$\underline{X}^t\,\underline{y} = \underline{X}^t\,\underline{X}\,\underline{\beta}$$

Note that this is not really least squares estimation as there is no redundancy.

$$\underline{\beta} = (\underline{X}^t\,\underline{X})^{-1}\,\underline{X}^t\,\underline{y}$$

In our example,

$$\begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{pmatrix} = \begin{pmatrix} 0.29 \\ 0.14 \\ 0.40 \\ 0.19 \end{pmatrix}$$

$$\underline{X}^t \underline{X} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$$(\underline{X}^t \underline{X})^{-1} = \begin{pmatrix} ¼ & 0 & 0 & 0 \\ 0 & ¼ & 0 & 0 \\ 0 & 0 & ¼ & 0 \\ 0 & 0 & 0 & ¼ \end{pmatrix}$$

$$\underline{X}^t \underline{y} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0.29 \\ 0.14 \\ 0.40 \\ 0.19 \end{pmatrix} = \begin{pmatrix} 1.02 \\ 0.16 \\ -0.36 \\ -0.06 \end{pmatrix}$$

$$\underline{\beta} = (\underline{X}^t \underline{X})^{-1} \underline{X}^t \underline{y} = \begin{pmatrix} 0.255 \\ 0.040 \\ -0.090 \\ -0.015 \end{pmatrix}$$

or,

$$y = 0.255 + 0.04\,T_{coded} - 0.09\,P_{coded} - 0.015\,T_{coded}\,P_{coded}$$

# What if we have three factors?

| $X_1$ | $X_2$ | $X_3$ |
|---|---|---|
| | | |
| -1 | -1 | -1 |
| -1 | -1 | 1 |
| -1 | 1 | -1 |
| -1 | 1 | 1 |
| 1 | -1 | -1 |
| 1 | -1 | 1 |
| 1 | 1 | -1 |
| 1 | 1 | 1 |

# Three Independent Variables

- Our example could be the partial oxidation of o-xylene
  - $x_1$ could be coded temperature
  - $x_2$ could be coded partial pressure of oxygen in the feed to the reactor
  - $x_3$ could be the coded presence or absence of bismuth molybdate catalyst

In two-level factorial designs, can incorporate Boolean variables

i.e. $x_3 = -1$ if no catalyst used

$= +1$ if catalyst used.

With three variables, the most general polynomial model that can be generated from a full 2 level factorial design is

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$$

Note: there are still no squared terms.

# Exercise (left to student)

Show that, for a full 2 level design in three variables, the $(\underline{X}^t \underline{X})^{-1}$ matrix is:

$$\begin{pmatrix} 1/8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & \\ & & & & & & & \\ 0 & 0 & . & . & . & & 0 & 1/8 \end{pmatrix}$$

# Fractional Factorial Designs

- The word “full” has been used a lot here. Are there 2 level designs that are not “full”?

- Consider the 3 variable design above. In a full design, $2^3$ or 8 runs are needed. What if we can only do 4 runs -- is there an intelligent way to choose the runs?

| $X_1$ | $X_2$ | $X_3$ |
|---|---|---|
| | | |
| -1 | -1 | -1 |
| -1 | -1 | 1 |
| -1 | 1 | -1 |
| -1 | 1 | 1 |
| 1 | -1 | -1 |
| 1 | -1 | 1 |
| 1 | 1 | -1 |
| 1 | 1 | 1 |

## (poor choice)

We could just take the first 4 runs (where $x_1$ is always –1), but then $x_1$ doesn't vary. Can we pick the 4 runs so that all of the $x_i$ terms vary?

## (good choice)

| $x_1$ | $x_2$ | $x_3$ |
|---|---|---|
| -1 | -1 | 1 |
| -1 | 1 | -1 |
| 1 | -1 | -1 |
| 1 | 1 | 1 |

Note: each $x_i$ varies and no columns are identical

However, you can confirm that

$x_1 = x_2 x_3$

$x_2 = x_1 x_3$

$x_3 = x_1 x_2$

These are termed "confounding" relationships.

and $x_1 x_2 x_3 = 1$

Thus, the 8 parameter model cannot be determined by this data (we should have known that you can't estimate 8 parameters with 4 runs).

Often, make the assumption that interaction effects are negligible (relative to linear effects). Then, fit a simpler model to the data:

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

but recognize that, for example, $\beta_2 x_2$ really represents $\beta_2 x_2 + \beta_{13} x_1 x_3$, and so on.

How did I come up with this subset of 8 runs that ensured that the linear effects wouldn't be confounded with each other?

# Generators

- I used the relationship $1 = x_1x_2x_3$ (i.e. the three way interaction term is set to the constant, or $\beta_0$ and $\beta_{123}x_1x_2x_3$ will be indistinguishable).
- Multiply the equation above by $x_1$:

$$x_1 = x_1^2x_2x_3$$

But $x_1^2 = 1$ in a 2 level factorial design, so

$$x_1 = x_2x_3.$$

Similarly,

$$x_2 = x_1 x_3$$

$$x_3 = x_1 x_2$$

Exercise (left to students)

Show that $x_1x_2x_3 = -1$ picks out the other four runs in the full $2^3$ design.

# Resolution

- In our 4 run subset of the $2^3$ design, (termed a $2^{3-1}$ fractional factorial design),
  - zero order terms (the constant) were confounded with third order terms
  - first order terms were confounded with second order terms
- Note that the sum of the orders of the confounded terms is three

Our 4 run subset of the $2^3$ design is then termed to have "resolution III" and is denoted as a

$2^{3-1}_{III}$ design

The higher the resolution, the better. If we took our "poor choice" subset where $x_1 = -1$, that would be resolution I.

# Plackett Burman Designs

- In general terms, fractional 2 level factorial designs are used to screen a large number of possibly influential variables to see if they truly influence the response. We usually assume that the effects of second order and higher terms (interactions) are much less than those of the constant and linear terms.

# Plackett Burman Designs

- The most efficient of these “screening” designs is due to Plackett and Burman*
- 12 runs can test up to 11 independent variables (other PB designs have 20, 24, 28, and so on for the number of runs)

* R.L Plackett and J.P. Burman, “The Design of Optimal Multifactorial Experiments”, Biometrika, 33, 305-325, (1946)

# Plackett Burman Designs

- The Plackett Burman designs are based on certain properties of Hadamard matrices
- All PB designs are of resolution III
- PB designs are sometimes termed “saturated”

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | + | + | + | + | + | + | + | + | + | + | + |
| 2 | - | + | - | + | + | + | - | - | - | + | - |
| 3 | - | - | + | - | + | + | + | - | - | - | + |
| 4 | + | - | - | + | - | + | + | + | - | - | - |
| 5 | + | + | - | - | + | - | + | + | + | - | - |
| 6 | - | + | + | - | - | + | - | + | + | + | - |
| 7 | - | - | + | + | - | - | + | - | + | + | + |
| 8 | + | - | - | + | + | - | - | + | - | + | + |
| 9 | + | + | - | - | + | + | - | - | + | - | + |
| 10 | + | + | + | - | - | + | + | - | - | + | - |
| 11 | - | + | + | + | - | - | + | + | - | - | + |
| 12 | + | - | + | + | + | - | - | + | + | - | - |

Plackett Burman design to test 11 variables in 12 runs

# Significance, Variances and Error Bounds

- The fitted models we have discussed utilize all of the experimental data to generate coefficients -- there are no “degrees of freedom” left to estimate variances
- If there is an independent estimate of the “pure error variance” (call it $s^2$) and if it can be assumed that this does not vary across the domain of interest, then

$$\underline{V}(\underline{\beta}) = (\underline{X}^t \underline{X})^{-1} \underline{s}^2$$

and $$\underline{V}(\underline{y}) = \underline{X} (\underline{X}^t \underline{X})^{-1} \underline{X}^t \underline{s}^2$$

# Note

Although the V( ) notation suggests that we are talking about variances, it should be obvious that we are only talking about estimates of variances. If we wish to find confidence intervals for the true variances, we will need to use the $\chi^2$ test; confidence intervals for the true values of the $\beta$ parameters or the predicted y values will require a t-test.

# Note to the Note

If the confidence interval for one of the β terms includes zero, the data are telling us that that particular combination of variables is not influencing the result (or if it does have an influence, the influence is less that the effects of the uncontrollable variables (i.e. “the noise”)). The parameter and the corresponding combination of variables should be omitted from the model.

# Aside on Nomenclature

If the prediction variance (V(y)) is only a function of the distance from the centre of the design (and not otherwise a function of position), we say that the design is rotatable. It can be shown that this leads to robust estimation of the response surface.

# Response Surface Designs

- Once we have determined which variables influence a response, we would like to quantify “how much” or determine a maximum or minimum over the domain under study
- 1.  If we have a really good idea of the functional relationship between dependent and independent variables, use a “parameter

estimation design". Briefly, test in areas of the domain where a small change in the parameters will yield a large change in the response. We will not deal further with this here.

2. If we do not know the functional form of the relationship, we may proceed by assuming an empirical relationship. The simplest functional form that will allow for local maxima and minima is the general quadratic:

1 variable: $y = \beta_0 + \beta_1 x_1 + \beta_{11} x_1^2$

2 variables: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2$

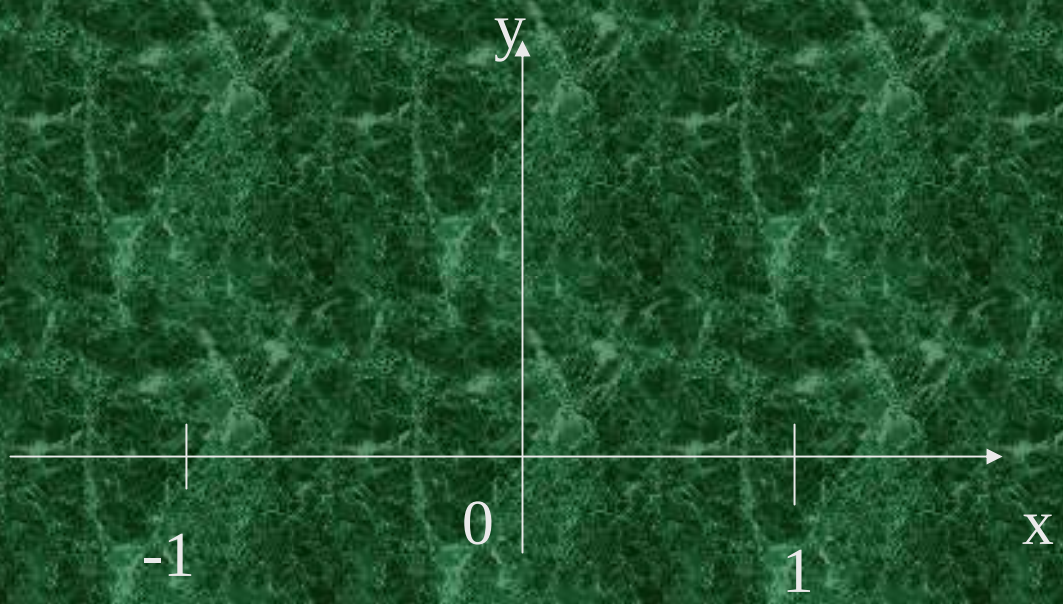
3 variables: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2$

and so on.

We cannot use 2 level factorial designs to estimate quadratic terms (recall that in 2 level designs, $x_i^2 = 1$). Therefore, we must use 3 levels (at least) for each independent variable.

# Three Level Factorial Designs

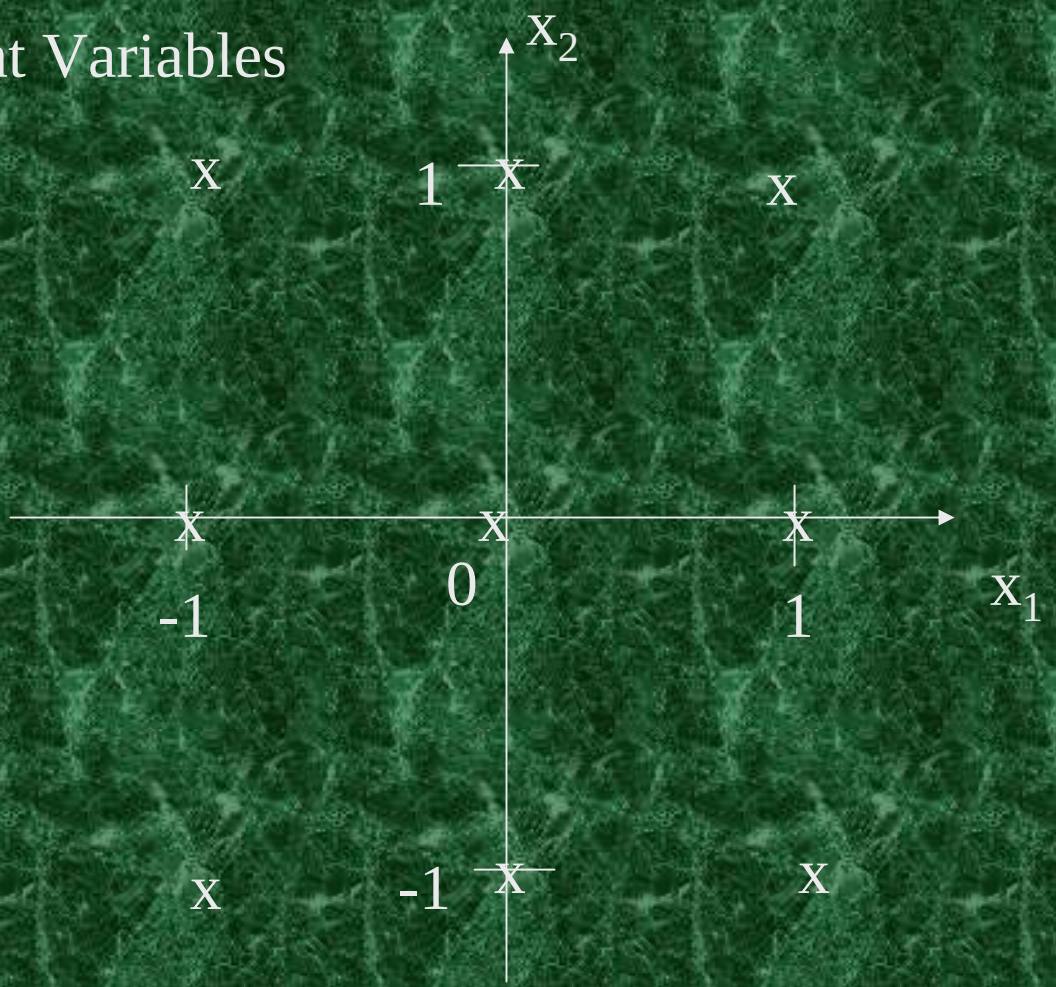
1 Independent Variable:




Test at $x_{coded}$ = -1, 0 and +1.

# Three Level Factorial Designs

2 Independent Variables



# Three Level Factorial Designs

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | -1 | -1 | -1 | 0 | 0 | 0 | 1 | 1 | 1 |
| $x_2$ | -1 | 0 | 1 | -1 | 0 | 1 | -1 | 0 | 1 |

Bonus: we only have to estimate 6 terms ($\beta_0$, $\beta_1$, $\beta_2$, $\beta_{12}$, $\beta_{11}$, $\beta_{22}$) using these 9 data points, so that an estimate of the variance can be made at the same time.

# Example: More on v for $NH_3$

Estimate a general quadratic model for the specific volume of ammonia with temperature and pressure as the independent variables, valid in the region from 0 to 100 $^oC$ and 100 to 300 kPa. Experimental data are available from a $3^2$ factorial design; these data are given in the following table.

# Specific Volumes of Ammonia in $m^3/kg$

| | 0°C | 50°C | 100°C |
|---|---|---|---|
| 100 kPa | 1.3145 | 1.5664 | 1.8145 |
| 200 kPa | 0.6471 | 0.7774 | 0.9035 |
| 300 kPa | 0.4243 | 0.5143 | 0.5997 |

# Other Response Surface Designs

- As the number of independent variables increases, the number of runs required for a full $3^n$ design greatly increases. For instance, if n = 3, $3^n = 27$; if n = 4, $3^n = 81$. If experiments cost time and/or money, this may be unacceptable.
- However, we do not really need this large number of runs. Consider

If we have 2 independent variables, we need to estimate 6 coefficients.

If we have 3 independent variables, we need to estimate 10 coefficients.

If we have 4 independent variables, we need to estimate 15 coefficients

In general, if we have n independent variables, we need to estimate (n+2)(n+1)/2 coefficients.

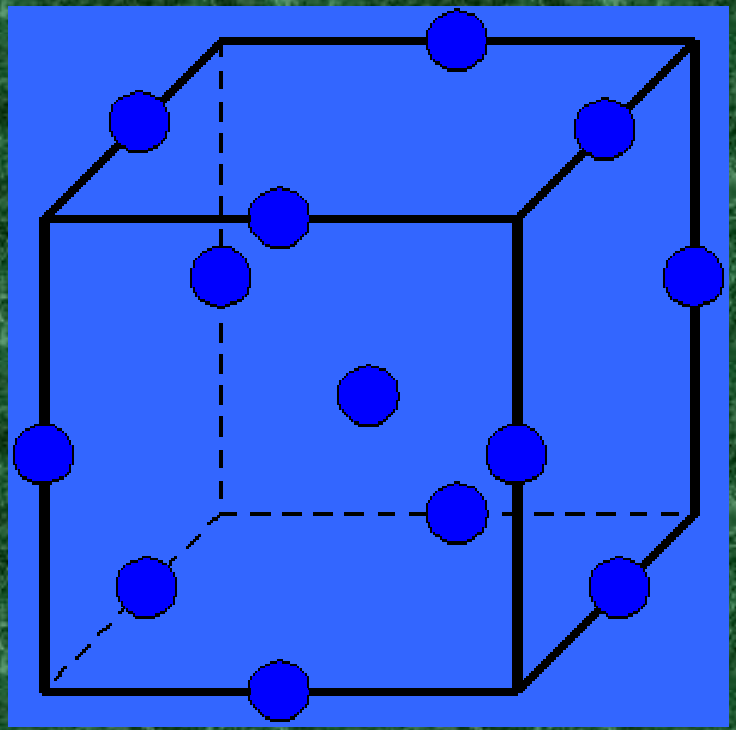
Obviously, $3^n$ runs may be overkill. Are there designs which still have all variables tested at at least three levels, but which have fewer runs?

# Box Behnken Designs

- The most widely used three level factorial design is due to Box and Behnken*
- Designs are available for three or more independent variables.
- Let us examine the three variable case.

* G.E.P. Box and D.W. Behnken, ”Some new three level designs for the study of quantitative variables”, Technometrics, 2, 455, (1960)

# 3 Factor Box Behnken Design



Note: there are 12 individual runs on the centre points of the edges of a cube and 3 points at the centre of the cube.

# 3 Factor Box Behnken Design

The actual runs are given by

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $x_2$ | -1 | -1 | 1 | 1 | 0 | 0 | 0 | 0 | -1 | 1 | -1 | 1 | 0 | 0 | 0 |
| $x_3$ | 0 | 0 | 0 | 0 | -1 | -1 | 1 | 1 | -1 | -1 | 1 | 1 | 0 | 0 | 0 |

The actual runs would, of course, be conducted in a random order.

# Box Behnken Example

- Consider a 10 m x 10 m plot of land from which a leaking gasoline tank had been removed. The contaminated soil had been returned to the excavation and clean fill added on top to compensate for the volume of the tank.
- The tank had been resting on a layer of compacted clay 4.8 m below grade. The clean fill is estimated to occupy the upper 2.5 m of the subsurface.

Use a Box Behnken design to estimate the parameters in a quadratic expression for hydrocarbon concentration.

If z is the distance below grade and x and y are rectangular co-ordinates relative to the centre of the 10 m x 10 m square,

$x = -5$ implies $x_{coded} = -1$ and $x = 5$ implies $x_{coded} = 1$

$y = -5$ implies $y_{coded} = -1$ and $y = 5$ implies $y_{coded} = 1$

$z = 5$ implies $z_{coded} = 1$, $z = 2.5$ implies $z_{coded} = -1$ and $z = 3.75$ implies $z_{coded} = 0$

Samples were taken and the following were the results:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| x | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 |
| y | -1 | -1 | 1 | 1 | 0 | 0 | 0 | 0 |
| z | 0 | 0 | 0 | 0 | -1 | -1 | 1 | 1 |
| $c_{HC}$ | 800 | 50 | 550 | 10 | 40 | 0 | 70 | 10 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| x | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| y | -1 | 1 | -1 | 1 | 0 | 0 | 0 |
| z | -1 | -1 | 1 | 1 | 0 | 0 | 0 |
| $c_{HC}$ | 10 | 0 | 60 | 5 | 1100 | 1050 | 1100 |

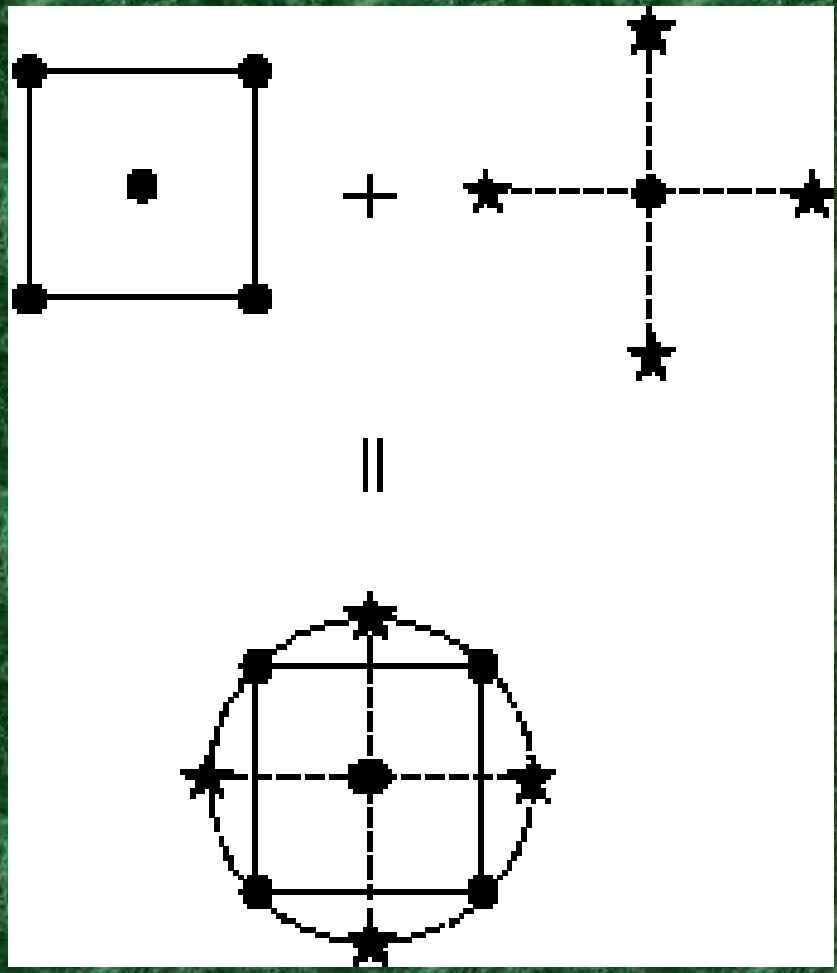
Note that hydrocarbon concentrations are given in mg/kg (ppm)

# Question for Discussion

Why isn’t there a Box Behnken design for two independent variables?

# Other Response Surface Designs

- Central composite designs (CCDs) are also used, although these tend to be less efficient than Box Behnken
- CCDs require five levels and usually consist of a 2 level design plus a “star-shaped” design. These were dealt with in the 1951 Box Wilson paper (qv).



# Construction of a CCD in two factors

CEI Inc.

# Mixtures and Scheffé Designs

- One area where one can encounter problems is in the optimal design of mixtures, because the physical requirement that the sum of all the mass fractions must be one reduces the rank of the X matrix
- The way around this and the way to approach experimental design is due to Scheffé*

*H. Scheffé, "Experiments with mixtures", J. Roy. Statist. Soc, B20, 344-360 (1958), and H. Scheffé "The simplex-centroid design for experiments with mixtures", J. Roy. Statist. Soc., B25, 235- 263 (1963)

# Henry Scheffé

CEI Inc.

In order to not have this problem with rank, Scheffé found that you could eliminate the constant.

?

Consider a linear function of the two components of a mixture, $x_1$ and $x_2$

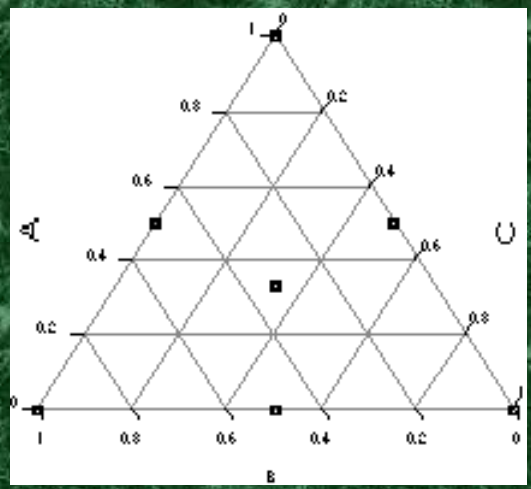
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Since $1 = x_1 + x_2$, we can write

$$y = \beta_0(x_1 + x_2) + \beta_1 x_1 + \beta_2 x_2$$

$$= (\beta_0 + \beta_1)x_1 + (\beta_0 + \beta_2)x_2$$

Scheffé also came up with simplex-centroid experimental designs. For three components,




The design points correspond to all permutations of the pure components (e.g., 1 0 0; 0 1 0; 0 0 1), the permutations of the binary blends (½ ½ 0; ½ 0 ½; 0 ½ ½) and the one blend involving equal parts of all three components.

The points in the simplex centroid design for three components are:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x_1$ | 1 | 0 | 0 | ½ | ½ | 0 | 1/3 |
| $x_2$ | 0 | 1 | 0 | ½ | 0 | ½ | 1/3 |
| $x_3$ | 0 | 0 | 1 | 0 | ½ | ½ | 1/3 |

Note that this design only contains 7 runs. A full quadratic model for three components requires 9, since we will eliminate the constant. To get sufficient data, it is usual to add

| | | | |
|---|---|---|---|
| $x_1$ | 2/3 | 1/6 | 1/6 |
| $x_2$ | 1/6 | 2/3 | 1/6 |
| $x_3$ | 1/6 | 1/6 | 2/3 |

Let’s look at an example that takes us back to experimental design’s agricultural roots. Synthetic mixtures of sand, silt and clay are used to grow potatoes under controlled conditions. The response is measured as kg potatoes yield per square metre of bed. Find the general quadratic model that best fits the data:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sand Fraction | 1 | 0 | 0 | ½ | ½ | 0 | 1/3 | 2/3 | 1/6 | 1/6 |
| Silt Fraction | 0 | 1 | 0 | ½ | 0 | ½ | 1/3 | 1/6 | 2/3 | 1/6 |
| Clay Fraction | 0 | 0 | 1 | 0 | ½ | ½ | 1/3 | 1/6 | 1/6 | 2/3 |
| Yield ($kg/m^2$) | 18 | 12 | 6 | 16 | 11 | 10 | 15 | 19 | 15 | 9 |

# Sequential Designs

A sequential design is one in which future experimental conditions depend on the results of previously conducted experiments. The one sequential design that we have seen is the central composite design. Another sequential design that is also used is the procedure known as Evolutionary Operation (or EVOP)

# Evolutionary Operation (EVOP)

The term Evolutionary Operation (or EVOP) is due to George Box*, although the method as presently known (using the "simplex" algorithm) is due to Spendley et al**; the numerical algorithm was further refined by Nelder and Mead***.

*G.E.P. Box, "Evolutionary operation: A method for increasing industrial productivity", Applied Statistics 6, 81-101 (1957).

**W. Spendley, G.R. Hext and F.R. Himsworth, "Sequential application of simplex designs in optimization and EVOP", Technometrics 4, 441-461 (1962).

***J.A. Nelder and R. Mead, "A simplex method for function minimization", Computer Journal 7, 308-313 (1965).

# Evolutionary Operation

The basic idea is to determine (by some other technique) the factors that influence the response. EVOP can then be used to determine what values of these factors to use to maximize (or minimize) the response.

If n factors have been found that influence the response, n + 1 runs that are linearly independent are made (e.g. if there are 2 important factors, 3 runs are made at conditions that do not lie on a straight line).

# Evolutionary Operation, cont'd

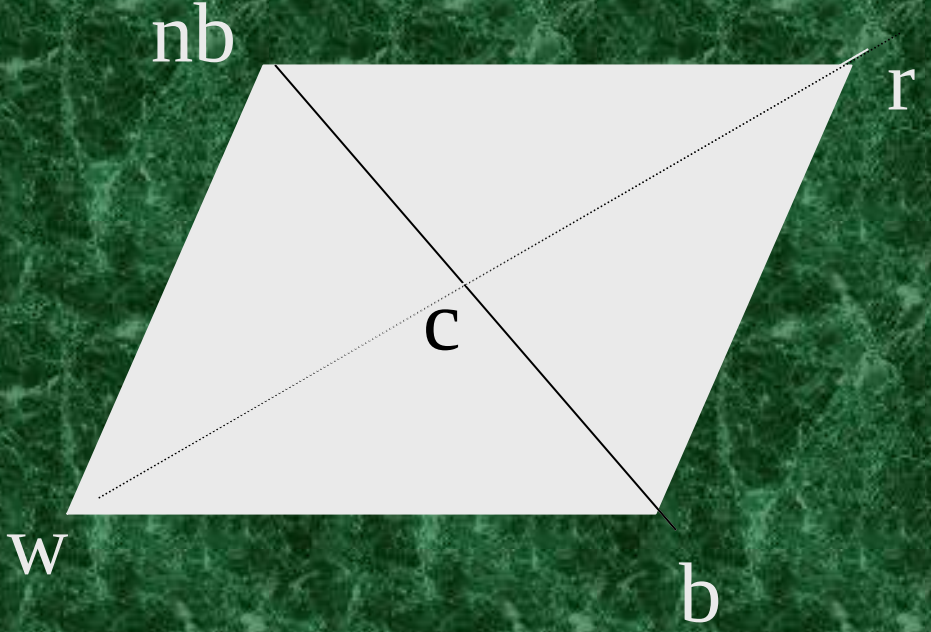
The n + 1 points in n-space form a simplex (3 points in 2-space form a triangle, 4 points in 3-space form a tetrahedron,…). The simplex method consists of discarding the worst result and testing at one new condition; this is done repeatedly until a maximum (or minimum) is reached.

# Evolutionary Operation, cont'd

One of the main advantages of EVOP is that after the process begins, the runs can often be made without major disturbances of the process under study. As the process continues, the changes in the input variables required by the process become smaller and smaller.

To illustrate the method, let us proceed with an obvious example with two input variables.

# Geometry of 2 – D Simplex Method




w worst point

b best point

nb next best

c centroid of points left after w removed

r reflection of w through c

# EVOP Example

Z is the response to two input variables x and y; Z is given by

$$Z = 100 / (1 + x^2 + y^2) \ .$$

Clearly, we can see that Z is maximized for $x = y = 0$ and takes on no other local maxima or minima for any finite, real x and y values.

Assuming that we do not know about this maximum, the simplex method tells us to find Z at 3 points:

| | | | |
|---|---|---|---|
| x | 2 | 3 | 2 |
| y | 0 | 3 | 2 |
| Z | 20 | 5.26 | 11.1 |

Clearly (3,3) is w, (2,0) is b and (2,2) is nb. Removing w, the centroid of the remaining points is (2,1).

To reflect w through c, either

1. Add the vector sum of c – w to c, or
2. Simply compute r = 2c – w

In our case, r = 2(2,1) - (3,3) = (1,-1)

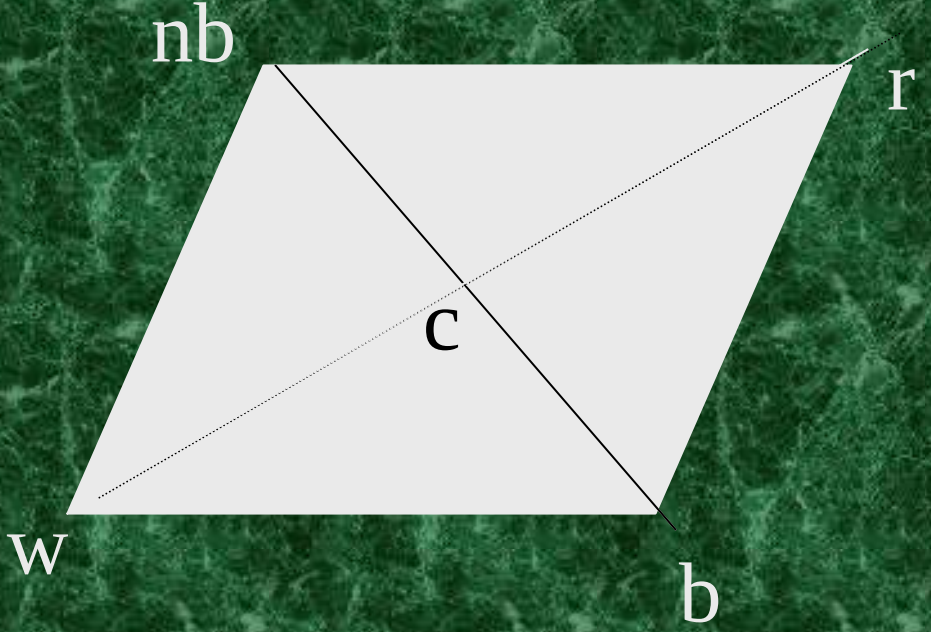
Computing Z(r) = 33.3, our three points in the simplex are now

| | | | |
|---|---|---|---|
| x | 2 | 2 | 1 |
| y | 0 | 2 | -1 |
| Z | 20 | 11.1 | 33.3 |

The worst point is at (2,2). The centroid after the worst point is removed is (3/2, -1/2). The reflection of the worst point through the centroid is 2c – w or (1,-3).

**Z(r) = 9.09 (worse than last worst)! Now what?**

# Geometry of 2 – D Simplex Method




w worst point

b best point

nb next best

c centroid of points left after w removed

r reflection of w through c

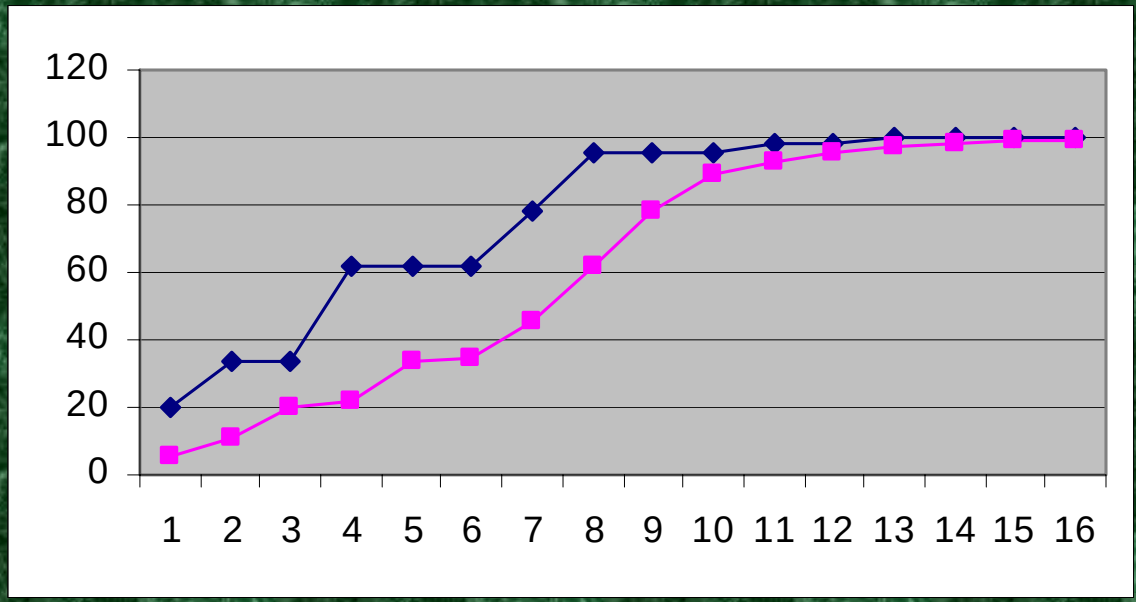
If the result at the reflection is worse than the result at w, replace r by the midpoint from w to c. In our case, this contraction would be ½{(2,2) + (3/2, -1/2)} or (1.75, 0.75). Z(1.75, 0.75) = 21.62. Now our three points entering the next iteration are:

| | | | |
|---|---|---|---|
| x | 2 | 1 | 1.75 |
| y | 0 | -1 | 0.75 |
| Z | 20 | 33.3 | 21.62 |

For the next iteration, w = (2, 0), c = (1.375, -0.125) and r = (0.75, -0.25).

Then, Z(r) = 61.5 and we are starting to converge on (0, 0) in this third iteration.

If we continue on with this process, we gradually approach the optimal conditions. Plotting the worst and best values of Z at each iteration gives the following graph.




Here the best value of Z at each iteration gives the upper curve and the worst value gives the lower curve. The "S-shaped" or logistic curve for the worst values is typical of many EVOPs.

Another observation from the graph is that, after 11 or 12 iterations, even the “worst” conditions are close to optimal. If this were a plant process, the EVOP process could continue as part of a continuous improvement plan making small incremental gains with little or no downside risk to the operation.

# Warning!

# Warning #1

The experimental runs used to converge on the maximum (or minimum) response are not good data to to use for generating a response surface model. If you need a model, generate data using a conventional response surface design.

# Warning #2

If there are uncontrollable inputs that are comparable in size to the effects of the controllable inputs, the EVOP process may never converge. The simplex may either oscillate with variations in the uncontrollable inputs or diverge completely.

# Warning #2, and Possible Cure

A possible cure is to deal with average responses, where the averaging is over a longer time period than the characteristic times associated with the uncontrollable inputs.

If the uncontrollable inputs have vastly different characteristic  times, this averaging may not be possible and a method other than EVOP must be used for optimization.

# Epilogue

With all the matrix algebra and n-space geometry, don’t forget that this all started out as a way to grow the best potatoes with the minimum amount of manure.