
# Design of Experiments

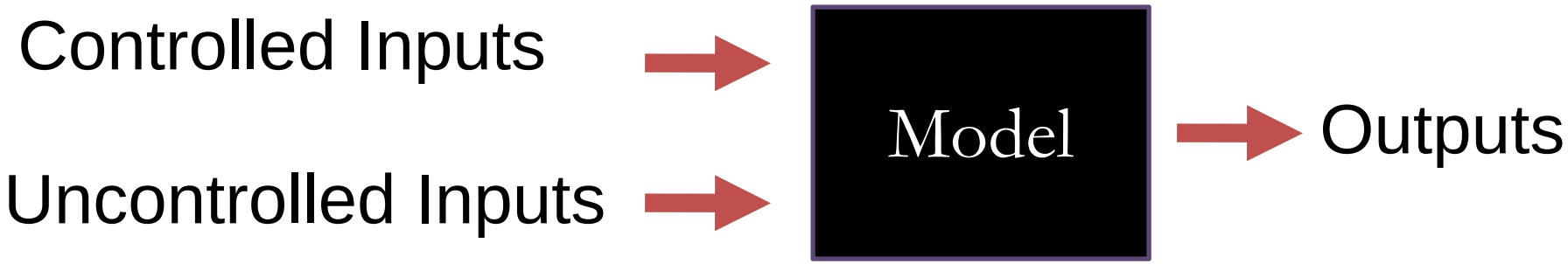
**Jeff Skinner, M.S.**

Biostatistics Specialist
Bioinformatics and Computational Biosciences Branch (BCBB)
NIH/NIAID/OD/OSMO/OCICB
http://bioinformatics.niaid.nih.gov
ScienceApps@niaid.nih.gov

# What Is a Designed Experiment?

- A designed experiment is planned, completed and analyzed using statistical considerations to increase efficiency
- Design of Experiments (DOE) has a rich history connected to agriculture, engineering and manufacturing
- Classic DOE attempted to coerce all experiments into a handful of well known, often inflexible, designs
- Modern statistics software packages allow non-experts to create flexible but sensible designs based on their own needs

# Testing A Model

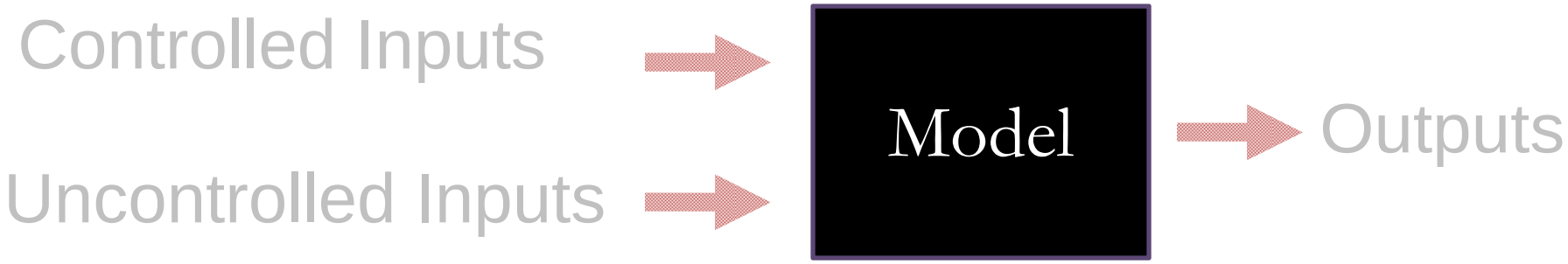



- Most scientific experiments can be visualized as a "black box" model of inputs and outputs
  - Effects of pH and temperature on protein binding yield
  - Effects of gene KO and microarray batch on gene expression

# Statistical Jargon

- Model inputs are **independent** variables
  - Controlled inputs are usually called ***factors***
  - Uncontrolled inputs are often called ***blocking variables***, ***covariates*** or ***nuisance variables***
  - All inputs may be called ***predictor variables***, e.g. **X**
- Model outputs are **dependent** variables
  - Outputs are also called ***response variables***, e.g. **Y**
- The "black box" is our statistical model
  - Linear models, nonlinear models, loglinear models, …

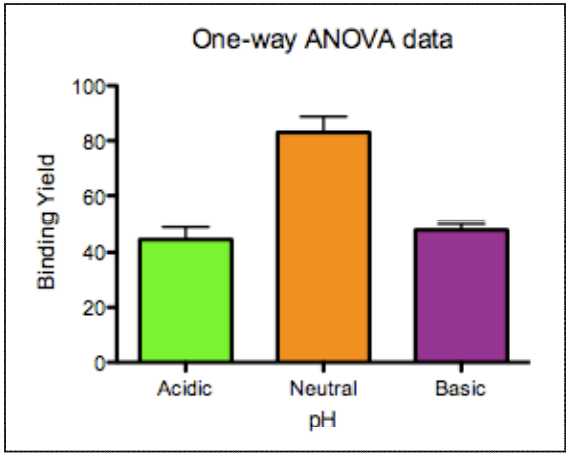
# Statistical Models

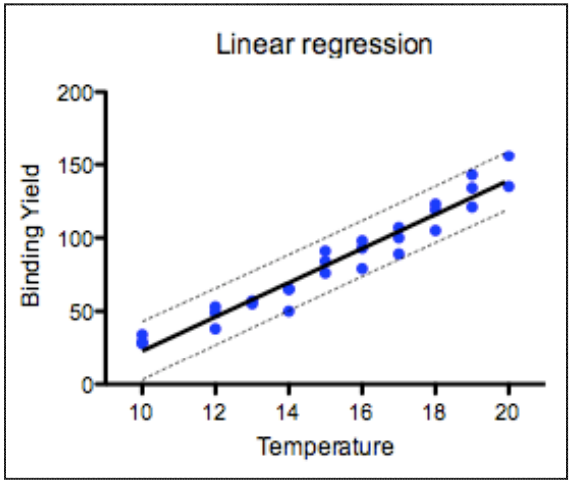


# Linear Models (continued)

- One-way Analysis of Variance (ANOVA) and simple linear regression are two familiar kinds of linear models

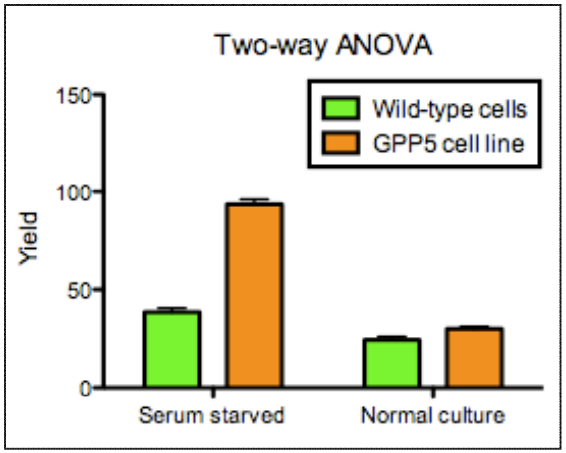
- ANOVA methods are used to compare mean response levels among groups
  - E.g. neutral pH binding yield is significantly higher than acidic or basic pH binding yield
- Regression explores linear relationships between predictor and response variables
  - E.g. for every one unit increase in temperature, mean binding yield increases by 11.62 units

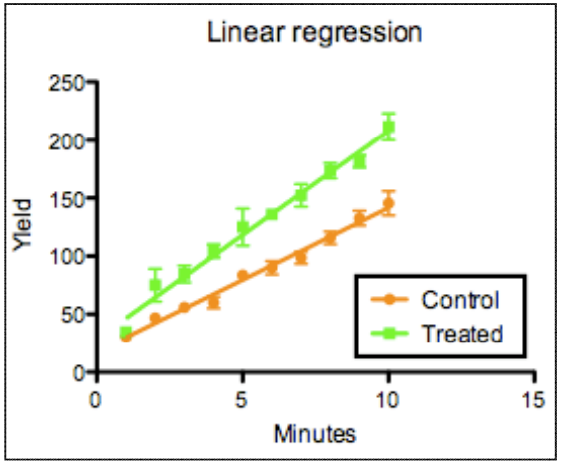


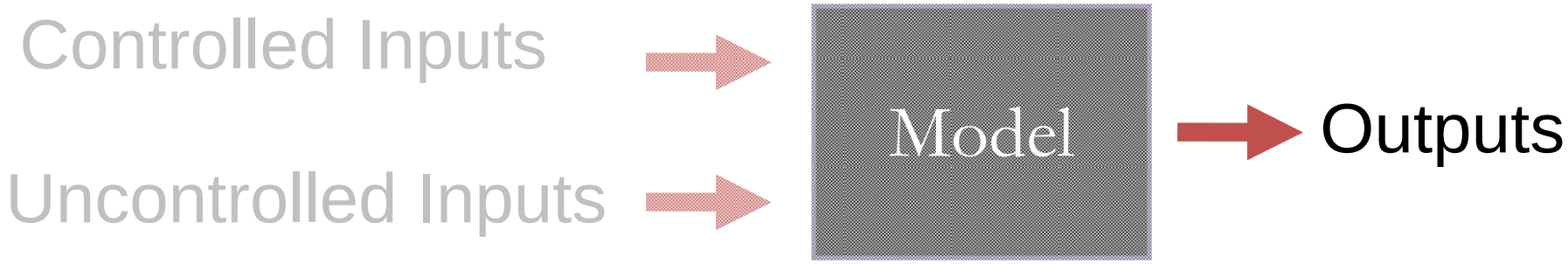
# Extensions of the Linear Model

- Multifactor ANOVA compares means among 2 or more factor effects (e.g. cell line and culture)
- Multiple regression looks at the association of 2 or more continuous predictors with the response
- Analysis of covariance (ANCOVA) includes both factor effect and continuous predictor variables



# Model Outputs



# Response Variables

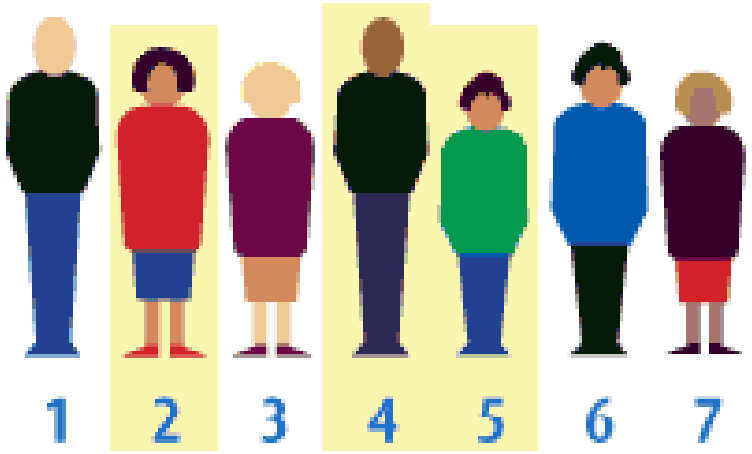
- Response variables determine the appropriate statistical model for the experiment
  - Most continuous responses use linear models (e.g. yield)
  - All categorical responses use generalized linear models
- Response measurements are taken from the units of sampled data in the experiment
  - Sample units may be selected from the population or assigned to the treatments in many different ways
  - Sample units may differ from experimental units

# Common Sampling Methods

- Simple Random Sampling (SRS)
  - Each sample in the population has an equal probability of being selected
- Stratified Sampling
  - The population is divided into fixed groups, or strata, and a SRS is selected from each of the strata
  - E.g. stratify a population by gender, then SRS both males and females
- Cluster Sampling
  - Several clusters are randomly selected, and all individuals are sampled
  - E.g. randomly choose 10 American cities and survey all citizens
- Systematic Sampling
  - Samples are chosen using an algorithm or heuristic
  - E.g. every fourth name from the phone book is chosen

# Simple Random Sampling

- Each unit in the population has an equal probability of being included in the sample
- Any differences among units will not be reflected in groups due to random sampling
- E.g. we don't want systematic differences between treatment groups to confound tests



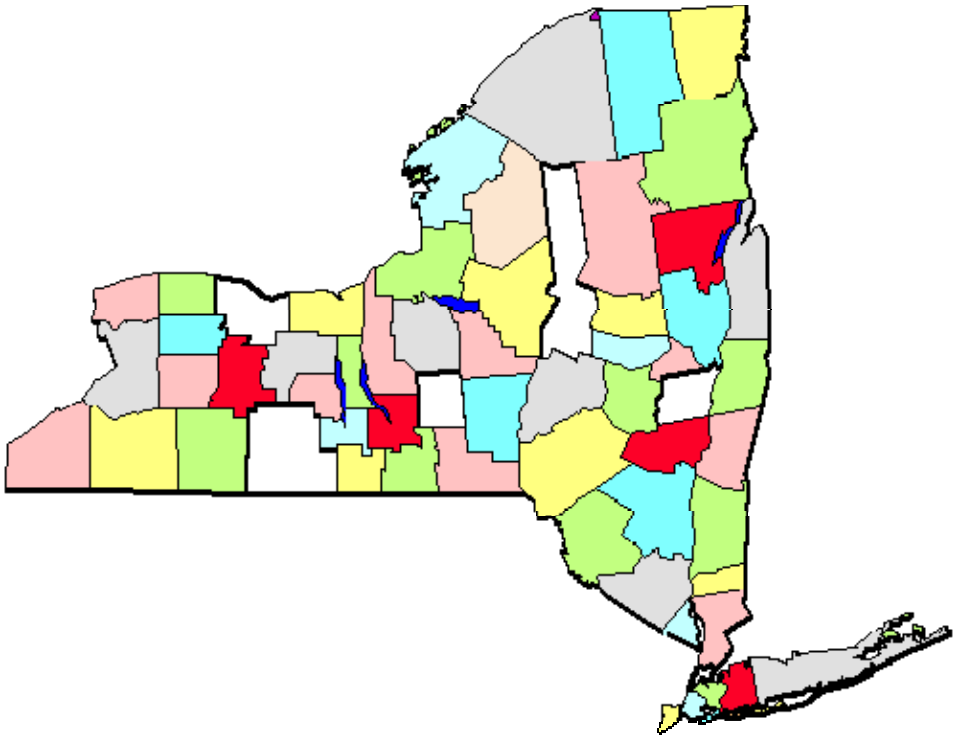

Assign Numbers, Auto-Generate Random Selections

# Stratified Sampling

- Want a representative sample of ALL lymphocytes in the human body
- Divide human body into several organ, tissue or cell types and SRS from each
  - E.g. T-cells, B-cells and NK-cells
  - E.g. Samples from lung, liver and gut tissue
- Ensures lymphocytes from all cell or tissue types are in the sample

# Cluster Sampling

- You want to sample city hospital data in NY state
- Randomly sample 5 of the 62 counties in NY
  - Each county is a “cluster”
- Collect all the relevant data from every hospital in the 5 sample counties



# Experimental Units vs. Sampling Units

- A ***treatment*** is a unique combination of all the factor levels from the controlled and uncontrolled inputs
- The ***experimental unit*** (EU) is the smallest entity that can receive or accept one treatment combination
- The ***sampling unit*** (SU) is the smallest entity that will be measured or observed in the experiment
- Experimental and sampling units are not always the same

# Example: EU and SU are the Same

- Suppose 20 patients have the common cold
  - 10 patients are randomly chosen to take a new drug
  - 10 patients are randomly chosen for the placebo
  - Duration of their symptoms (hours) is the response variable

- EU and SU are the same in this experiment
  - Drug and placebo treatments are applied to each patient
  - Each patient is sampled to record their duration of symptoms
  - Therefore EU = patient and SU = patient

# Example: EU and SU are different

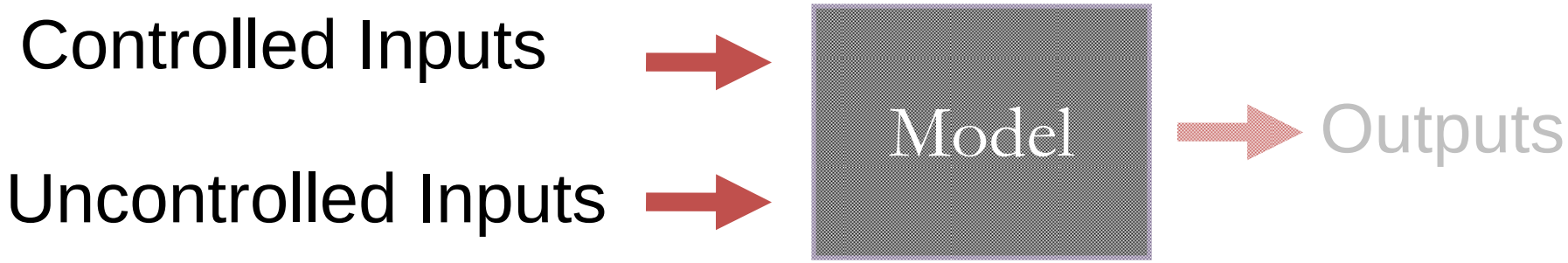
- 20 flowers are planted in individual pots
  - 10 flowers are randomly chosen to receive dry fertilizer pellets
  - 10 flowers are randomly chosen to receive liquid fertilizer
  - All six petals are harvested from each flower and petal length is measured as the response variable

- EU and SU are different in this experiment
  - Fertilizer treatment is applied to the individual plant or pot
  - Measurements are taken from individual flower petals
  - Therefore EU = plant and SU = petal **(pseudo-replication)**

# Pseudo-Replication

- Confusion between EU's and SU's can artificially inflate sample sizes and artificially decrease p-values
  - E.g. It is tempting to treat each flower petal as a unique sample (n = 6 x 20 = 120), but the petals are ***pseudo-replicates***
  - "Pseudoreplication and the Design of Ecological Field Experiments" (Hurlbert 1984, Ecological Monographs)
- Pooling samples can create pseudo-replication problems
  - E.g. 12 fruit flies are available for a microarray experiment, but must pool flies into 4 groups of 3 flies each to get enough RNA
  - Once data are pooled, it is not appropriate to analyze each individual separately in the statistical model

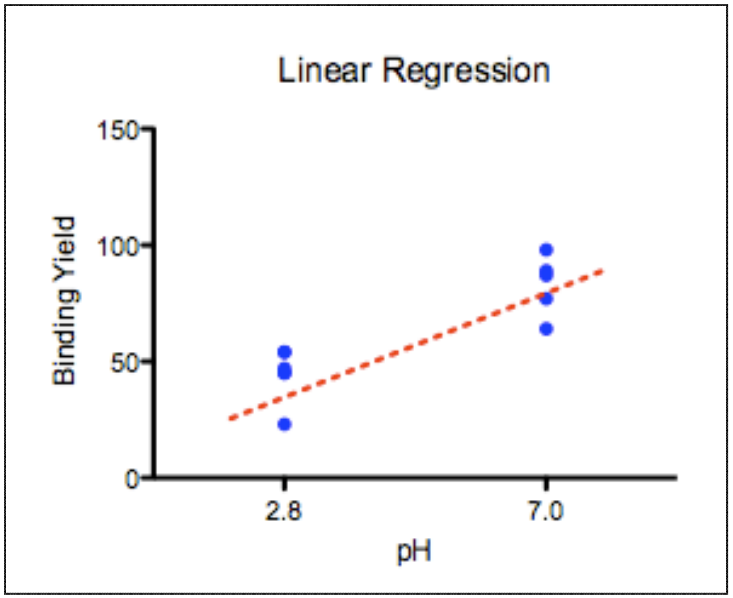
# Model Inputs

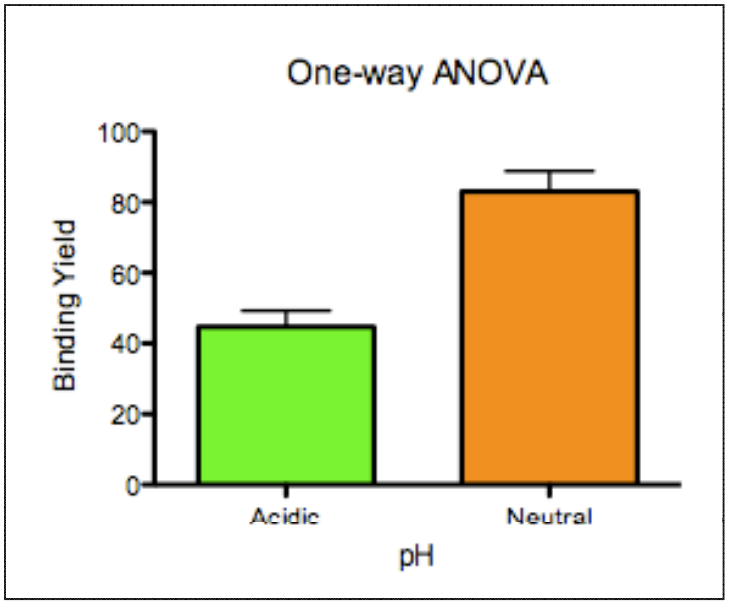


# Controlled Inputs

- Controlled inputs are the variables that most interest the experimenter
- Controlled inputs can be manipulated and randomly assigned by the researcher
- Controlled inputs can be treated as factor effects, even if the variable is continuous

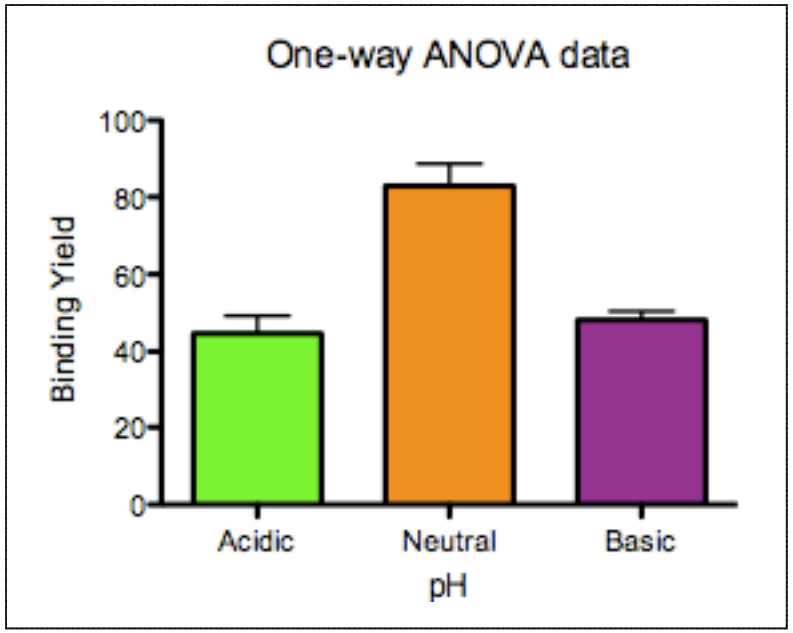
# Relationship Between Regression and ANOVA

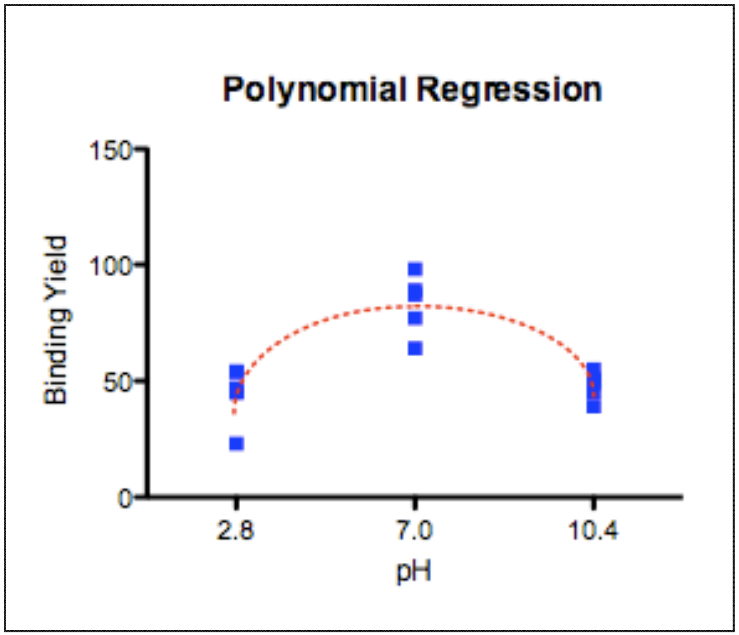



- Continuous variables can be treated as factors, if researchers only use 2-3 replicated values

# Curved Relationships




- You need at least 3 replicated **X** values or 3 factor levels to estimate a curved relationship

# Uncontrolled Inputs

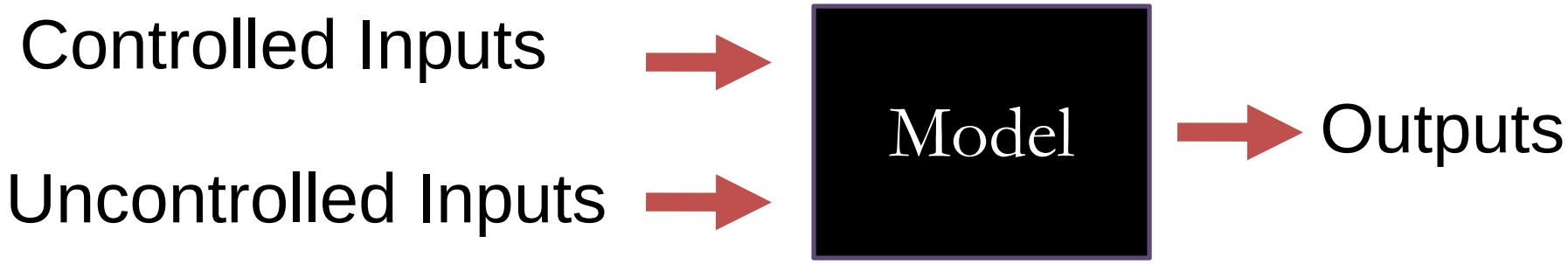
- Uncontrolled inputs are measurable traits of the sampling units that do not interest the researcher, but must be included in the statistical model

- Uncontrolled categorical variables are **nuisance factors**
  - E.g. gender, race, disease status, cell type, smoking status, …
- Uncontrolled continuous variables are **covariates**
  - E.g. body mass, age, calcium intake, cigarettes per week, …
- Uncontrolled discrete variables are **blocks**
  - E.g. agricultural plot, microarray chip, chemical batch, subject

# Blocking Variables

- Sometimes the phrase ***block variable*** is used to describe consecutive measurements from one factor level
  - E.g. We might block an experiment by the **gender** variable, if we collect all of the male and female samples separately
  - These variables may be better described as ***whole plot effects***
- The phrase ***block variable*** can also describe categorical or discrete variables consumed during the experiment
  - E.g. If **chemical batch** is a variable in an experiment, you will only get a fixed amount of experiment runs from each batch
  - **Subject** variables are often treated as blocks, because one subject can only accept one treatment at a time

# A Complete Picture of DOE



# Why Use a Designed Experiment?

- You want an unbiased experiment (**Basic Designs**)
- You want to know how many samples to collect in an upcoming experiment (**Power and Sample Size**)
- You know dozens of variables *might* affect a protein binding process and you want to test them with one reasonably-sized experiment (**Screening Designs**)
- You want to optimize a fermentation experiment by controlling its duration, substrate concentration and temperature to maximize yield, minimize costs and target a specific mass (**Response Surface Designs**)

# Completely Randomized Design

- The simplest type of designed experiment may be the completely randomized design (CRD)
- In the CRD, experimental units are randomly assigned to the factor level groups using simple random sampling
  - E.g. Any medical studies where all patients can be randomly assigned to drug or placebo groups might be a CRD
- Randomly assigning treatments to the SU's will eliminate biases from other correlated variables
  - E.g. biases from gender, age, weight or comorbidities in a medical study

# Randomized Complete Block Design (RCBD)

Four 30° C incubators with 6 trays each

Each incubator receives 3 trays from each drug

4 incubator blocks and 2 drug treatments

- Often CRD is impossible or inappropriate, so researchers must restrict randomization to control nuisance effects
  - E.g. Imagine a malaria experiment comparing the effects of two drugs on mosquitoes stored in one of four 30° C incubators with 6 trays each
- Randomized Complete Block Design (RCBD) experiments arrange samples into blocks by the nuisance factor(s)
  - E.g. Each incubator is a block and mosquitoes are randomly assigned to one of the two drug treatments and one of 3 trays in each of the 4 incubators

# Within and Among Block Variance

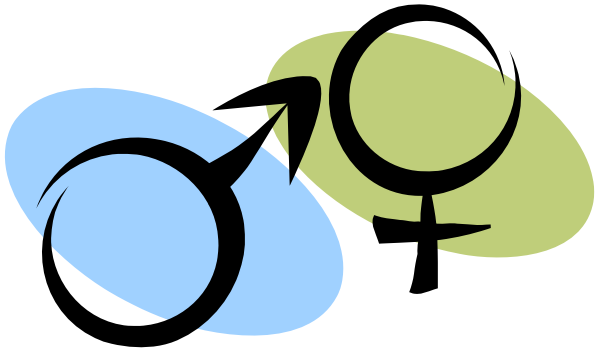
$$Y_{ij} = \theta + \alpha_i + \beta_j + e_{ij}$$

- $Y_{ij}$: parasite load
- $\theta$: intercept
- $\alpha_i$: among drug treatments
- $\beta_j$: among incubator block errors
- $e_{ij}$: within incubator block errors

- Samples from the same incubator should be correlated within one another, creating differences among incubator blocks
- RCBD experiments separate error within blocks from the error among blocks to increase statistical power
  - Differences among drug treatments are compared to the within block error
- CRD experiments ignore differences among blocks, confounding the errors within and among blocks for reduced statistical power
  - Among drug treatments are compared to within **and among** block error

# Random vs. Fixed Effects

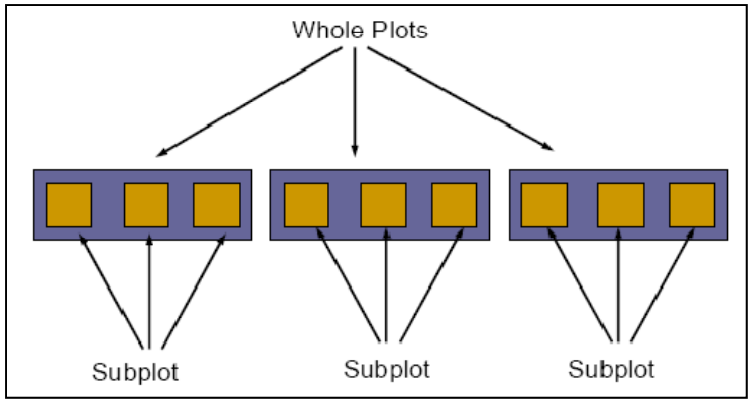
Subject effects are random



Gender effects are fixed

- Subject effects are random because the subjects in a experiment are a sample from the population of all possible subjects
- Gender effects are fixed because there are only two genders

# Split-plot Design




12 mice: 6 infected, 6 uninfected

3 infected males, 3 infected females, …

4 samples taken from each mouse

Each sample treated with one of 2 different drugs

Whole plot (mouse) EU's: Infection, gender

Subplot (sample) EU's: drug treatment

- Split-plot design experiments model experiments where whole plots and subplots represent different EUs
  - Whole plots are often locations, subjects, objects or factors that are difficult to change (e.g. temperature in an incubator)
  - Subplot effects are typically the effects of highest interest
  - Subplot effects are tested with higher power than whole plot

# Multiple Variances in Split-plot Design

$$Y_{ij} = \theta + \alpha_i + b_j + \gamma_k + e_{ijk}$$

| $Y_{ij}$ | $\theta$ | $\alpha_i$ | $b_j$ | $\gamma_k$ | $e_{ijk}$ |
|---|---|---|---|---|---|
| Response | intercept | whole plot treatments | whole plot error | subplot treatments | subplot errors |

- Whole plot effects are tested against the whole plot standard error
  - Whole plot standard error = $\sigma_e^2 + k\sigma_b^2$, where $\sigma_e^2$ is the subplot error, $\sigma_b^2$ is the whole plot error and k is the number of subplot treatments
  - Whole plot error component is often a random subject effect
  - Tests of whole plot effects have reduced power (larger standard error)
- Subplot effects are tested against the subplot standard error
  - Subplot error = $\sigma_e^2$ usually represents the random error between samples
  - Tests of the subplot effects have the most power (smaller standard error)

# Split-plot vs. RCBD

- Split-plot experiments deal with multiple EU's, while RCBD can have a single EU = SU
- Blocks in a RCBD experiment are usually nuisance variables that do not interest the experimenter
  - Blocks included to account for additional variation
  - Blocks are usually treated as random effects
- Whole-plot effects do interest experimenters
  - Whole-plot effects are often fixed effects

# Crossed vs. Nested Factors

| | | Water levels | |
|---|---|---|---|
| | | Low | High |
| Fertilizer | Low |  |  |
| | High |  |  |

crossed factors






nested factors

# Crossed and Nested Factors

- Two factors are crossed if all of their factor levels can occur together in a single experimental unit
  - E.g. Fertilizer level (high or low) and water level (high or low) are crossed because all combos are possible (LL, LH, HL, HH)
- Two factors are nested if the levels in one variable depend on the levels of the other variable for each EU
  - E.g. car manufacturer (Ford or Chevy) and car model (Mustang or F150, Corvette or Silvarado) are nested factors, because there is no Chevy F150
- Need to properly specify factors as crossed or nested
  - Cannot run nested variables as crossed effects in the model
  - Subplot effect is nested within the whole plot in a split-plot model

# Power and Sample Size

# How Many Samples?

- You are planning a biological experiment, but you do not know how many samples to collect
  - E.g. mouse experiments, microarray experiments, vaccine production
- One sample per group
  - Cannot computer errors, statistical tests or *p*-values
- Too few samples per group
  - Will I find any significant *p*-values?
  - Have I accurately represented my population?
- Too many samples per group
  - Did I waste time, money or other resources?
  - Are my statistical results biologically meaningful?

# Two Ideologies

**Karl Pearson**
1857-1936

Correlation coefficient
Linear regression
Chi-square test

**Sir Ronald A Fisher**
1890-1962

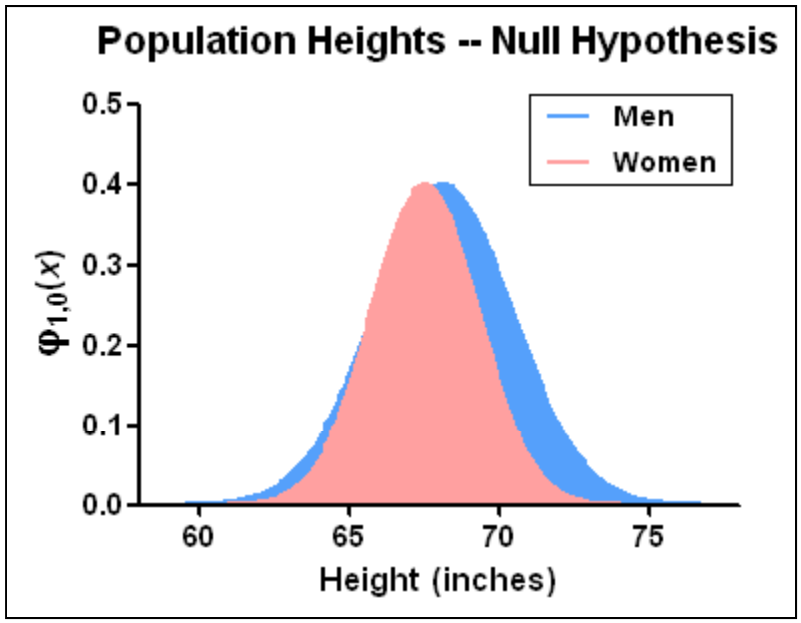
Analysis of Variance
Design of Experiments
Fisher's Exact Test

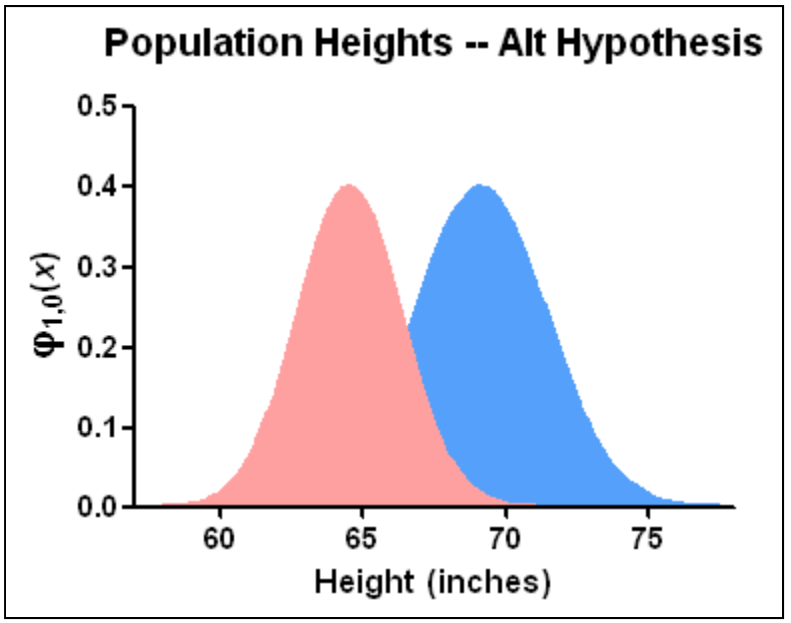
- Collect as many samples as possible (Karl Pearson)
  - More samples = more information about the population
- Collect a smaller representative sample (R.A. Fisher)
  - Detecting a significant result with a small sample produces stronger conclusions because it requires a larger effect size

# Recall the Statistical Testing Process

- Formulate null and alternative hypotheses
  - E.g. (null) $H_0$: $\mu_1 = \mu_2$ vs. (alternative) $H_A$: $\mu_1 \neq \mu_2$
- Calculate the appropriate test statistic
  - E.g. Student's t-test, linear regression, ANOVA F-test, …
- Compute the probability of observing the test statistic (i.e. your sample data) under the null hypothesis
  - I.e. Compute a p-value
- Make a statistical decision
  - Reject the null hypothesis or Fail to Reject the null hypothesis
- Make a biological conclusion
  - E.g. New drug reduces viral load, neutral pH produces highest yield, …

# Null and Alternative Hypotheses



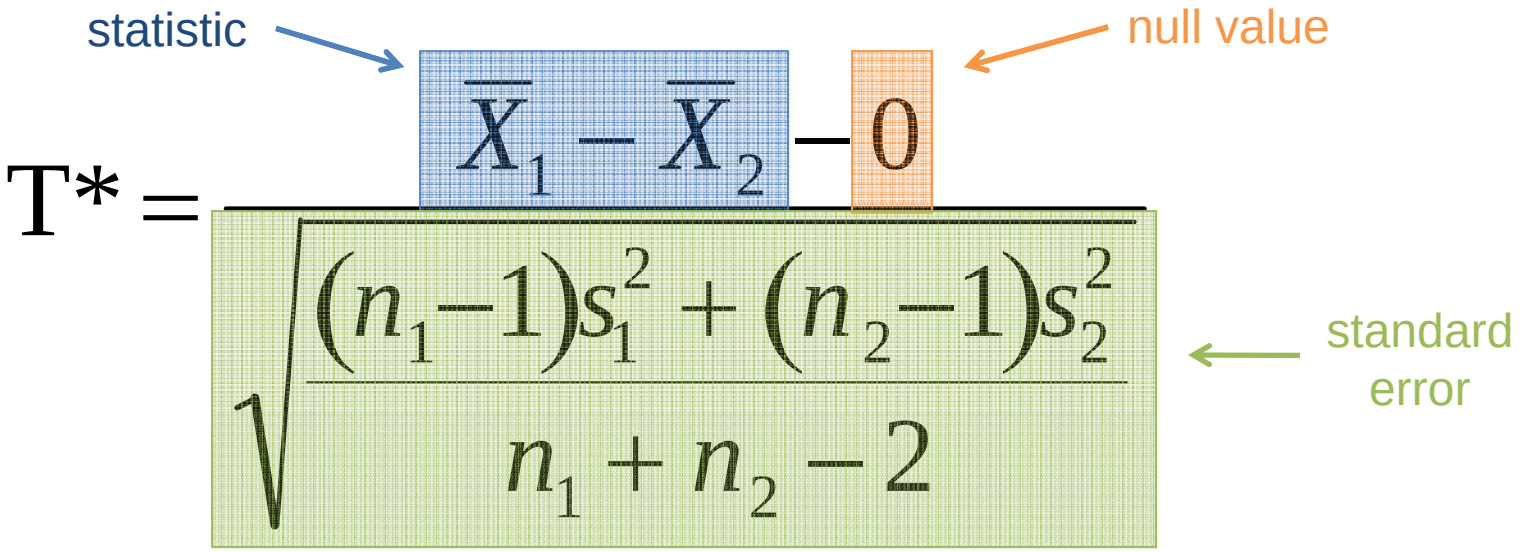

- Men and women are equal height vs. men taller than women
- (null) $H_0$: $\mu_M - \mu_W \leq 0$ vs. (alternative) $H_A$: $\mu_M - \mu_W > 0$

# What Is a Statistical Test?

$$\text{Test} = \frac{\text{Difference}}{\text{Error}} = \frac{\text{Statistic} - \text{Null Value}}{\text{Error}}$$

- Almost all tests used in inferential statistics can be generalized as the ratio of a "difference" over an "error"
  - Difference between a statistic and null value (usually 0)
  - A statistic is nothing more than a numeric summary of the experimental data with respect to the null hypothesis
  - Null value is an assumption about the population under the null hypothesis and error is estimate of the sampling distribution error

# Example: Two-sample Student's *T*-test



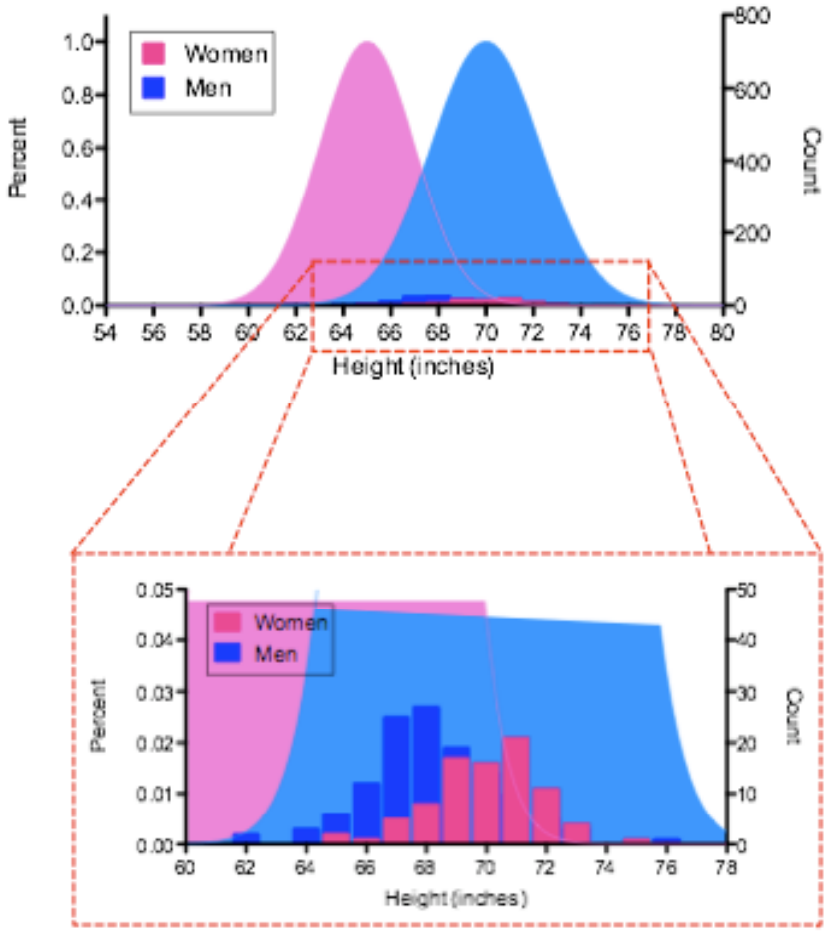

- The "statistic" in a two-sample *t*-test is a difference between the two sample means and the null value is zero
  - The hypothesis $\mu_1 = \mu_2$ implies $\mu_1 - \mu_2 = 0$
- The standard error is an estimate of the common variance

# Type I and Type II Errors

| | | Actual population difference? | |
|---|---|---|---|
| | | Yes | No |
| Was the difference detected by the statistical test? | Yes | OK | Type I Error (False Positive) |
| | No | Type II Error (False Negative) | OK |

- Different types of experiments attempt to minimize Type I errors, Type II errors or both kinds of errors
  - E.g. Type II errors are more important in medical testing

# Type I and Type II Errors - Example




- Suppose the average man is 5” taller than the average woman
  - Male population average 70”
  - Female population average 65”
- Samples from the population may not be representative
  - By chance we sample 120 former women’s basketball players
  - By chance we sample 120 men that are shorter than average
- Type I and Type II errors reflect the imperfections of sampling

# Power and Sample Size Analysis

- Power represents the probability of detecting an significant result whenever it truly occurs
- Statistical power is related to sample size and other characteristics of the experiment
- The goal is to determine the power achieved by a certain sample size or determine the sample size necessary to achieve the desired power

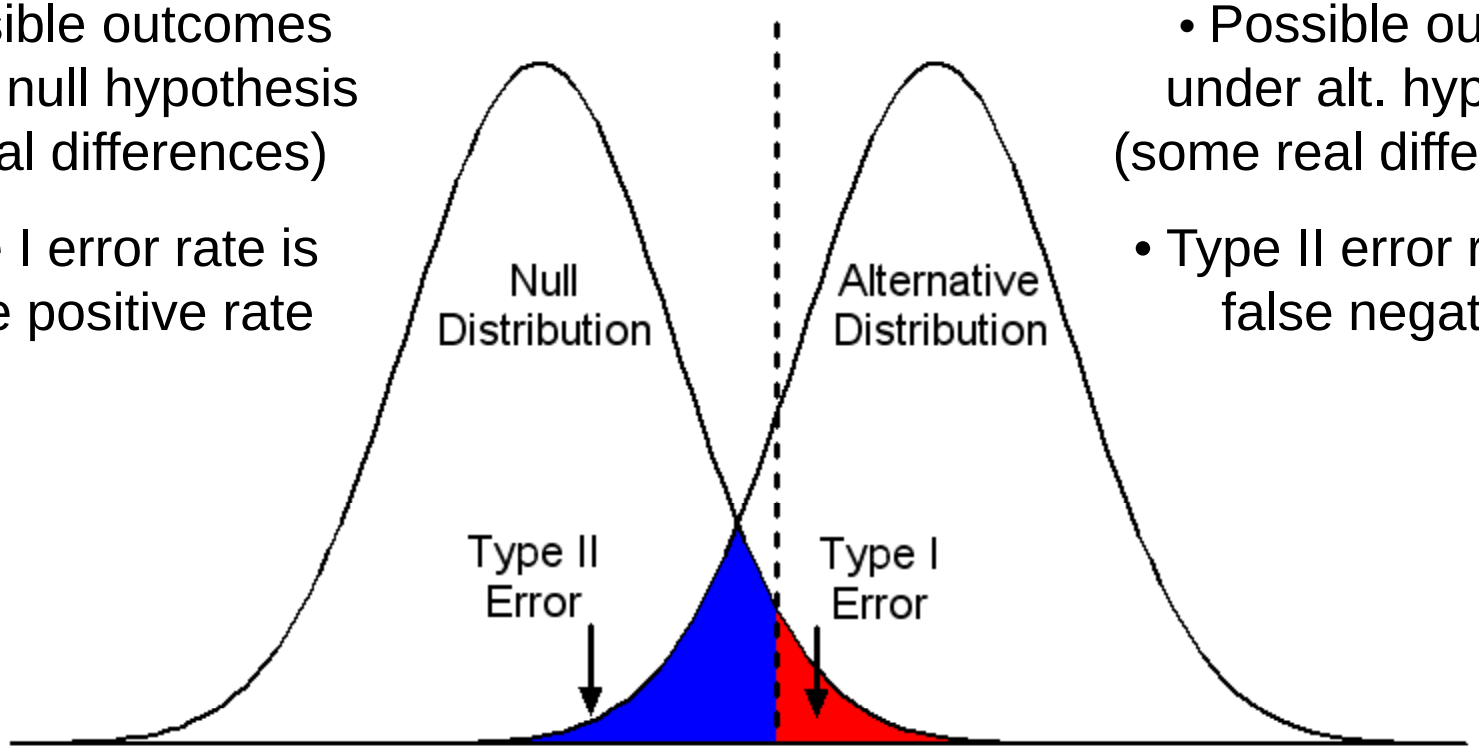
# Statistical Testing

Null distribution

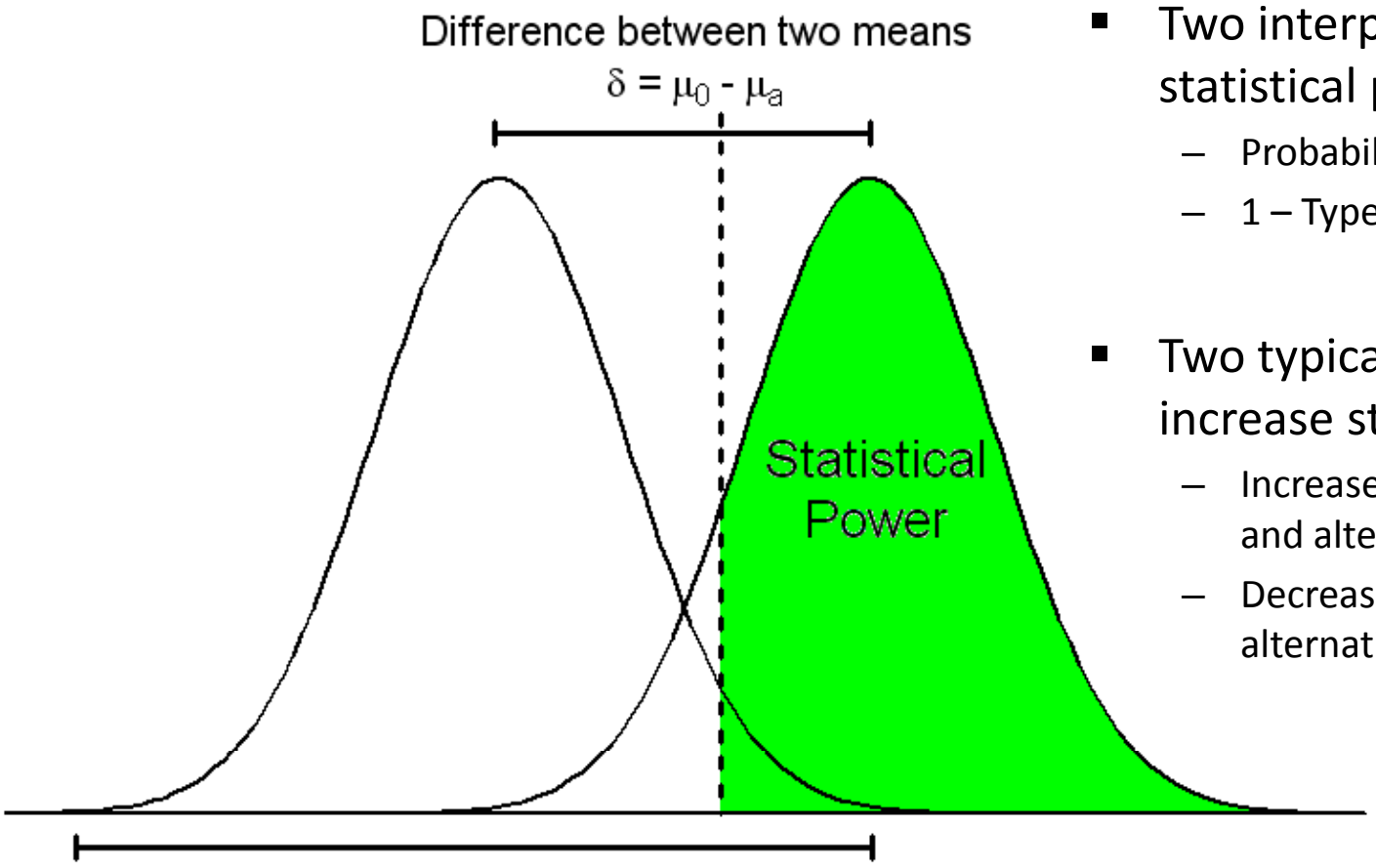
- Possible outcomes under null hypothesis (no real differences)
- Type I error rate is a false positive rate

Alt. distribution

- Possible outcomes under alt. hypothesis (some real differences)
- Type II error rate is a false negative rate

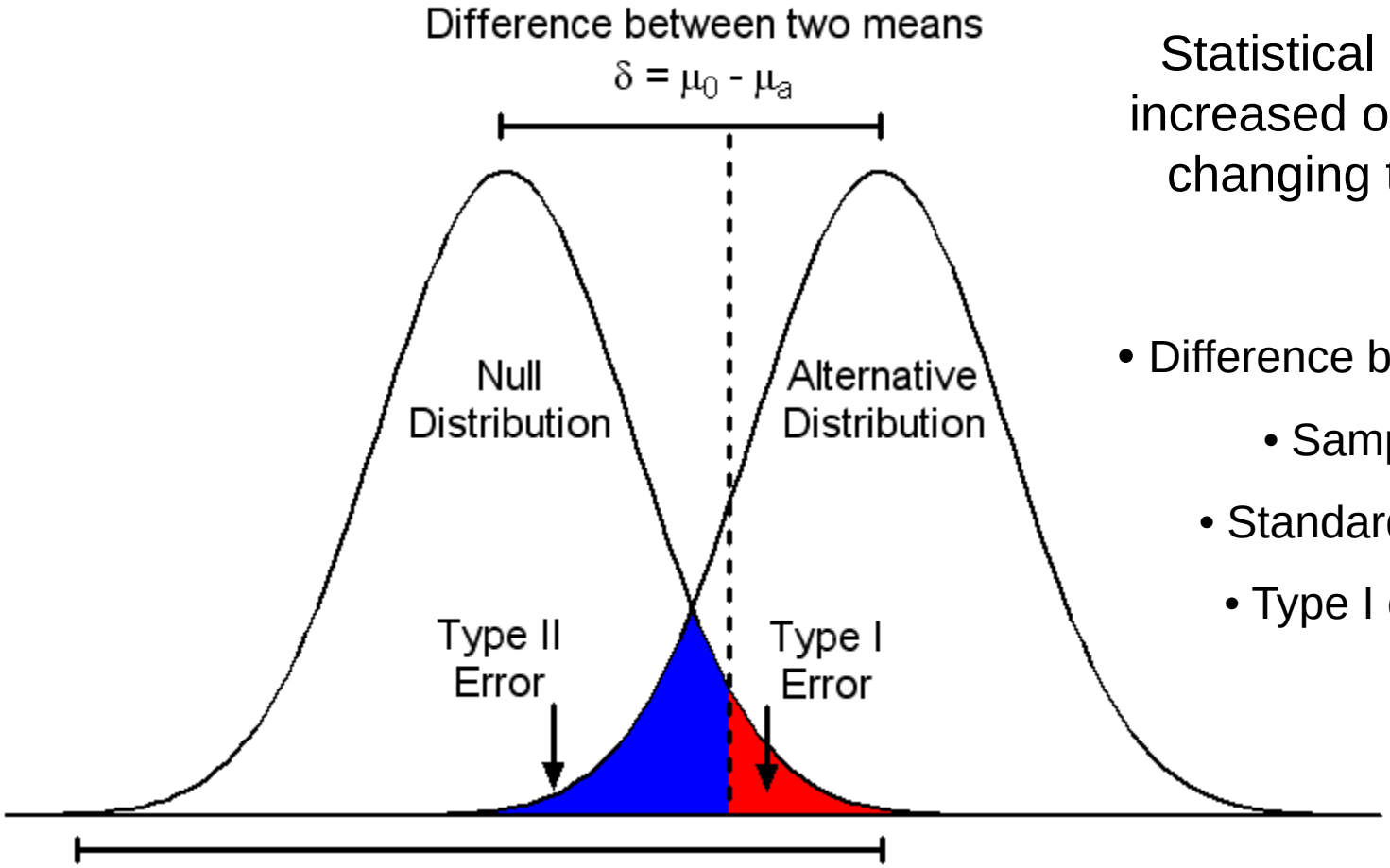


# Statistical Power




- Two interpretations of statistical power:
  - Probability of a true positive
  - 1 – Type II error rate
- Two typical methods to increase statistical power
  - Increase distance between null and alternative curves
  - Decrease the width of null and alternative curves

# Estimating Power and Sample Size




Standard error of the mean

$$SE = \frac{s}{\sqrt{n}}$$

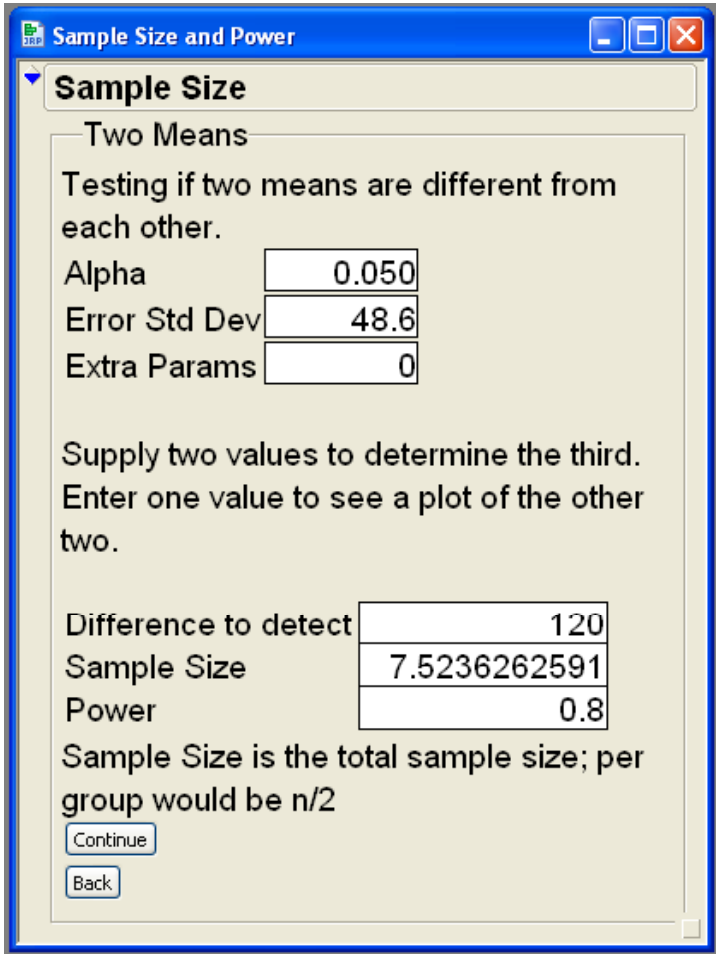
Statistical power can be increased or decreased by changing the values of:

- Difference between means, $\delta$
- Sample size, *n*
- Standard deviation, *s*
- Type I error rate, $\alpha$

Find a minimum sample size (*n*) by setting power, *s* , $\delta$ and $\alpha$ equal to constants from “pilot data”

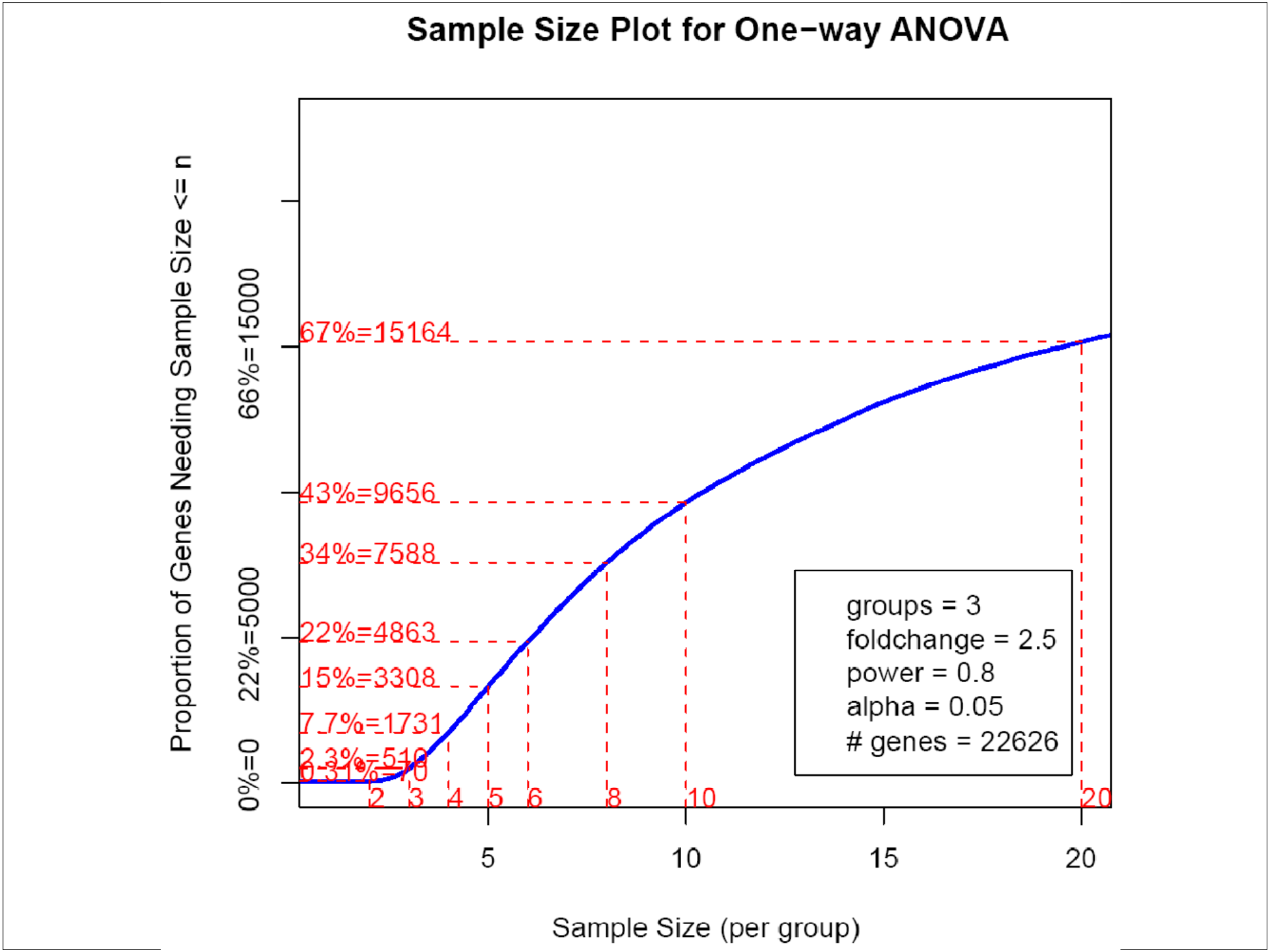
# Production Yield

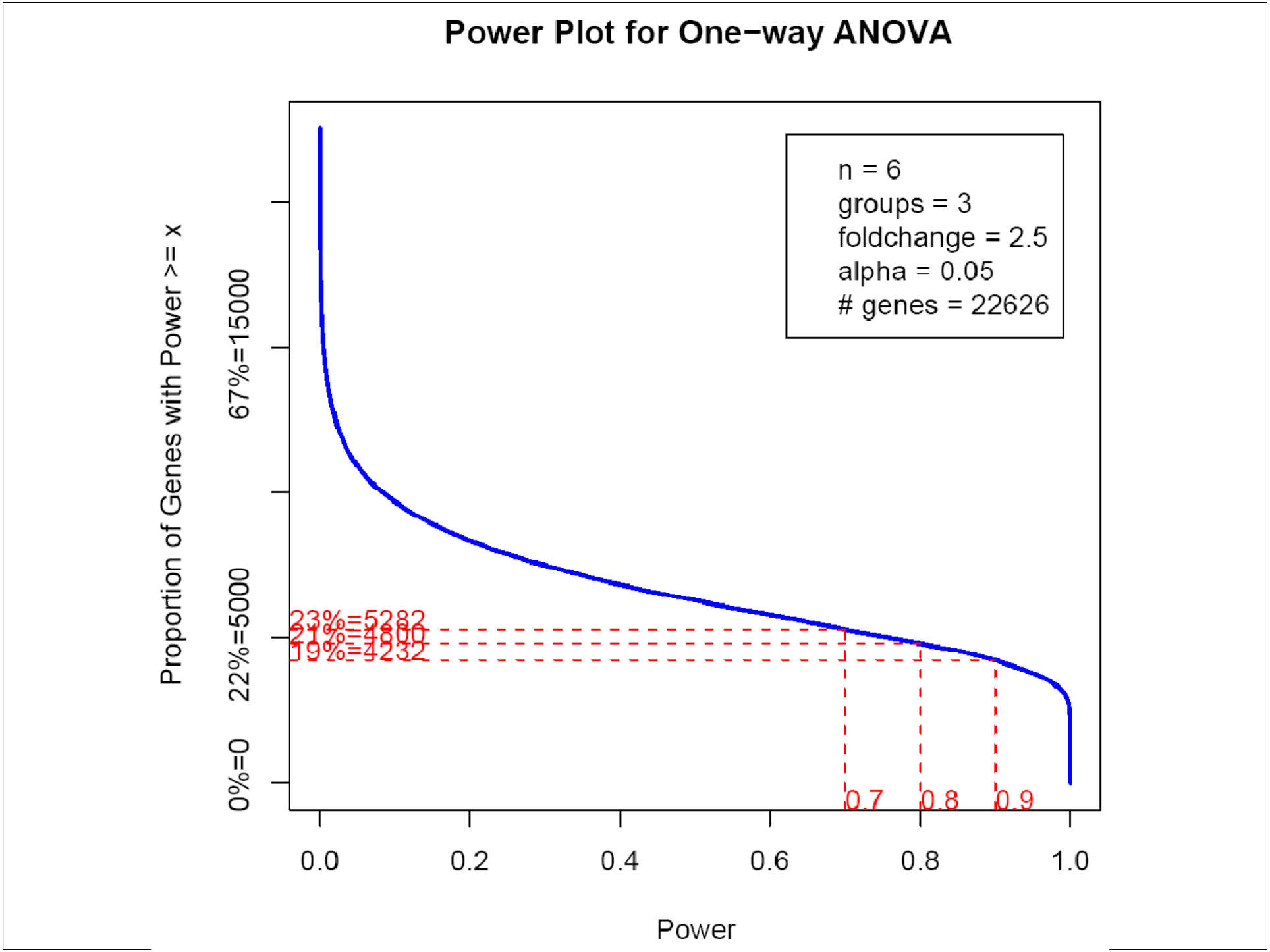
- Suppose you believe a new method should increase yield by 20% or about 120 grams
- You want 80% power to detect a 120 gram difference in yield with a significance level of 0.05 and a std dev of 48.6 grams
  - Need n = 8 samples total
  - Need n = 4 samples per group

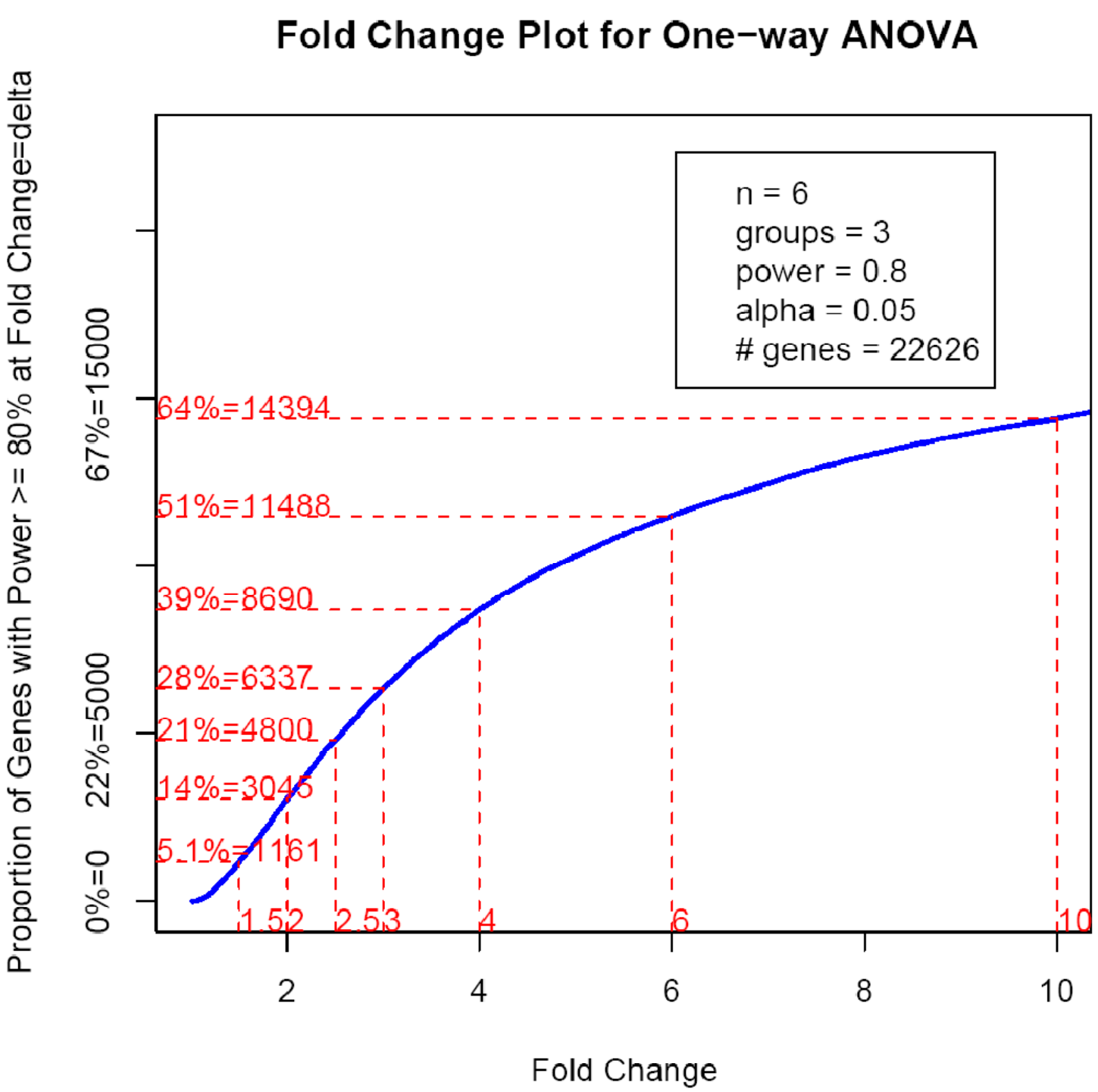


# Microarray Sample Size

- Microarray experiments test thousands of hypotheses, each with a unique variance and fold-change value
- Input variance estimates for each gene and apply the Bonferroni adjustment for multiple comparisons
- Plot the proportion of genes achieving the desired power, fold-change and sample size values


Sample Size Plot for One−way ANOVA
Proportion of Genes Needing Sample Size <= n
0%=0
22%=5000
66%=15000
67%=15164
43%=9656
34%=7588
22%=4863
15%=3308
7.7%=1731
2.3%=510
0.31%=70
2
3
4
5
6
8
10
20
5
10
15
20
Sample Size (per group)
groups = 3
foldchange = 2.5
power = 0.8
alpha = 0.05
# genes = 22626


Power Plot for One−way ANOVA
n = 6
groups = 3
foldchange = 2.5
alpha = 0.05
# genes = 22626
Proportion of Genes with Power >= x
0%=0
22%=5000
67%=15000
23%=5282
21%=4800
19%=4232
0.7
0.8
0.9
0.0
0.2
0.4
0.6
0.8
1.0
Power


Fold Change Plot for One−way ANOVA
n = 6
groups = 3
power = 0.8
alpha = 0.05
# genes = 22626
64%=14394
51%=11488
39%=8690
28%=6337
21%=4800
14%=3045
5.1%=1161
1.52
2.53
4
6
10
0%=0
22%=5000
67%=15000
Proportion of Genes with Power >= 80% at Fold Change=delta
2
4
6
8
10
Fold Change

# Why Use a Screening Experiment?

- One multi-factor experiment is more efficient than multiple experiments one-variable-at-a-time
  - Save time, money, materials, animals, …
  - A single multifactor experiment can utilize larger sample sizes per group than several smaller experiments
- Multi-factor experiments allow researchers to detect important interactions between variables
  - More accurate estimation of both main effects and interactions
  - Avoid bad interpretations due to Simpson's paradox

# A Simple Example

- You want to examine the effects of temperature, pH and substrate concentration on a protein binding yield
  - Each run requires 1 hour of bench time and costs $500
- Three experiments performed one-variable-at-a-time:
  - Three experiments with 10 runs each (30 runs total)
  - Sample size n = 5 per group (e.g. 5 observations with low pH)
  - Total cost $15,000 and 60 hours of bench time
- One designed multifactor experiment:
  - One experiment with 16 total runs (n = 8 per group)
  - Total cost $8000 and 32 hours of bench time

**Almost double the sample size for half the cost**

# A Simple Example

| Run | pH |
|---|---|
| 1 | 10 |
| 2 | 4 |
| 3 | 10 |
| 4 | 4 |
| 5 | 4 |
| 6 | 10 |
| 7 | 10 |
| 8 | 4 |
| 9 | 10 |
| 10 | 4 |

| Run | Temp |
|---|---|
| 1 | 15 |
| 2 | 25 |
| 3 | 15 |
| 4 | 25 |
| 5 | 15 |
| 6 | 25 |
| 7 | 15 |
| 8 | 25 |
| 9 | 15 |
| 10 | 25 |

| Run | Conc |
|---|---|
| 1 | 30 |
| 2 | 40 |
| 3 | 30 |
| 4 | 40 |
| 5 | 40 |
| 6 | 30 |
| 7 | 30 |
| 8 | 40 |
| 9 | 30 |
| 10 | 40 |

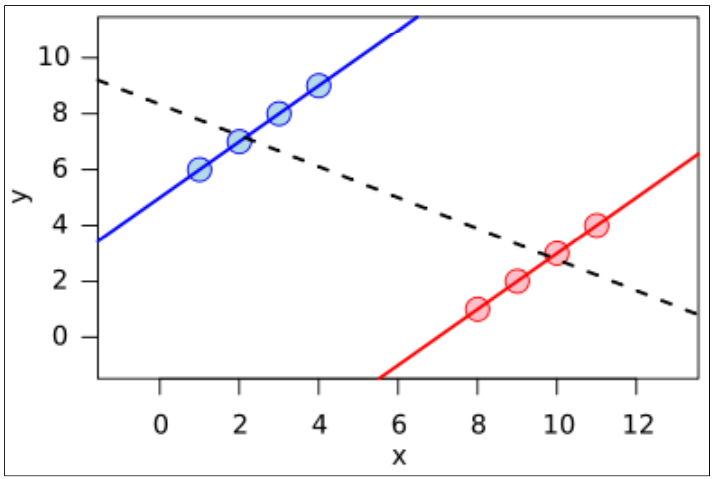
3 Experiments, one-variable-at-a-time

| Run | Temp | pH | Conc |
|---|---|---|---|
| 1 | 15 | 4 | 30 |
| 2 | 25 | 4 | 30 |
| 3 | 25 | 10 | 30 |
| 4 | 25 | 4 | 30 |
| 5 | 25 | 10 | 40 |
| 6 | 25 | 10 | 30 |
| 7 | 15 | 4 | 40 |
| 8 | 15 | 10 | 30 |
| ... | ... | ... | ... |
| 16 | 25 | 10 | 40 |

Designed Experiment

# Simpson's Paradox

- Do birth control pills lower the risk of heart attack?
  - Study finds risk of heart attack is 20% lower for women on birth control pill
  - Stratify women by age (< 35 years old) and there is no relationship between heart attack risk and the pill
- Simpson's Paradox
  - Relationship between two variables changes with the presence or absence of a third variable in the model
  - Often caused by confounding variables or differences in sample size




Relationship between androgen and estrogen levels in men and women

| | Tues | Wed | Total |
|---|---|---|---|
| Scott H. | 63/90 | 2/10 | 65/100 |
| Jeff S. | 9/10 | 31/90 | 40/100 |

Difference in **Call of Duty 4** shooting accuracy between two players

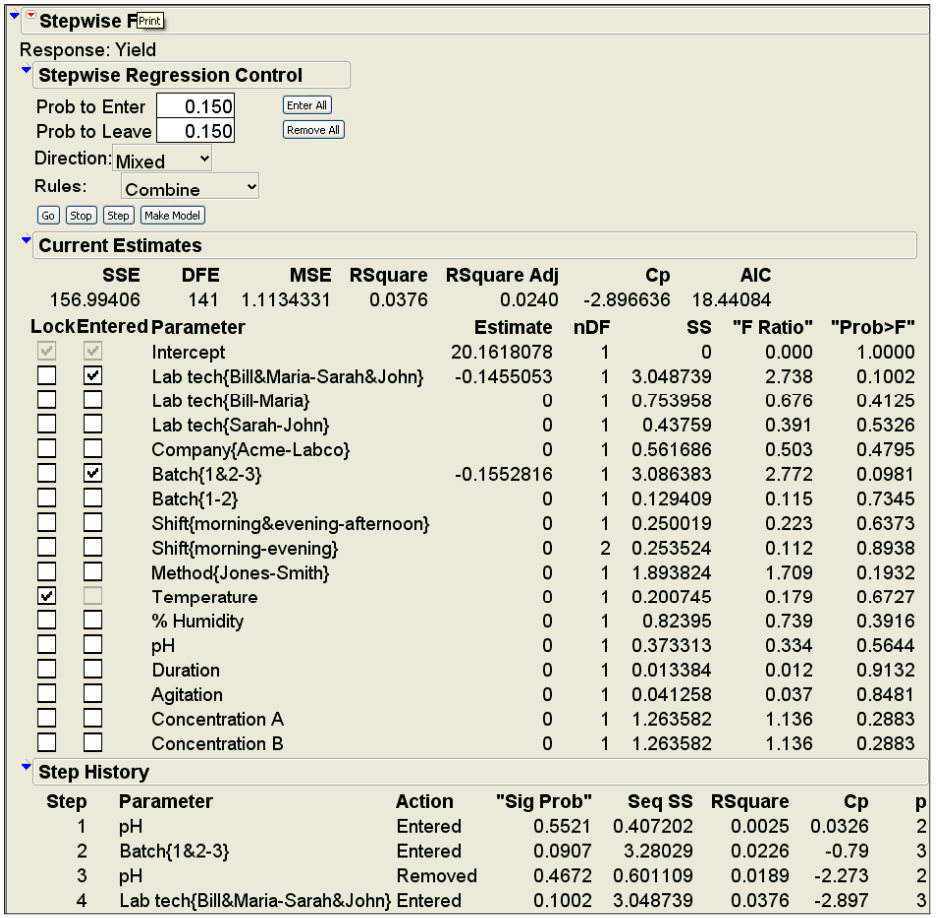
# Which Experimental Factors Most Affect Protein Binding?

- What do I use if I have already collected my data?
  - Stepwise regression methods
  - JMP Partitioning Platform

- What do I use if I am still planning my experiment?
  - Use traditional screening designs from a book
  - JMP DOE Screening Platform or Factorial Platform
  - JMP DOE Custom Design Platform

# Stepwise Regression Methods

- Use "Stepwise" personality from the Fit Model menu
- Specify the F-to-enter and F-to-remove values
- AIC and Mallow's $C_p$ are reported for best subsets selection methods
- Never interpret coefficients or predictions from a model selection procedure



Stepwise F Print

Response: Yield

Stepwise Regression Control

Prob to Enter 0.150 Enter All

Prob to Leave 0.150 Remove All

Direction: Mixed

Rules: Combine

Go Stop Step Make Model

Current Estimates

| SSE | DFE | MSE | RSquare | RSquare Adj | Cp | AIC |
|---|---|---|---|---|---|---|
| 156.99406 | 141 | 1.1134331 | 0.0376 | 0.0240 | -2.896636 | 18.44084 |

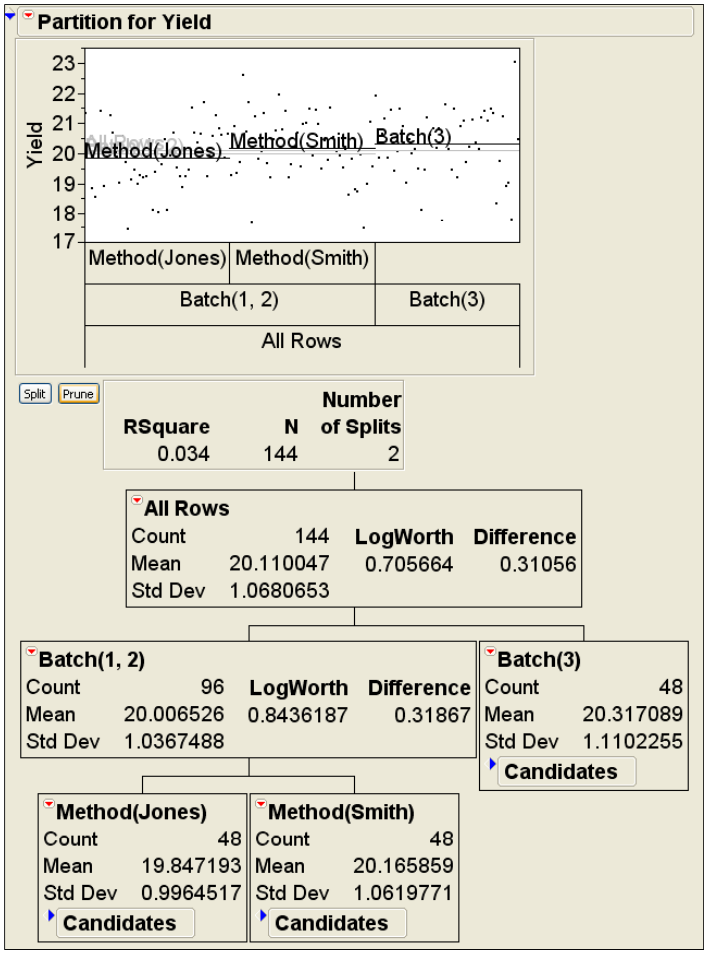
| Lock | Entered | Parameter | Estimate | nDF | SS | "F Ratio" | "Prob>F" |
|---|---|---|---|---|---|---|---|
| ☑ | ☑ | Intercept | 20.1618078 | 1 | 0 | 0.000 | 1.0000 |
| ☐ | ☑ | Lab tech{Bill&Maria-Sarah&John} | -0.1455053 | 1 | 3.048739 | 2.738 | 0.1002 |
| ☐ | ☐ | Lab tech{Bill-Maria} | 0 | 1 | 0.753958 | 0.676 | 0.4125 |
| ☐ | ☐ | Lab tech{Sarah-John} | 0 | 1 | 0.43759 | 0.391 | 0.5326 |
| ☐ | ☐ | Company{Acme-Labco} | 0 | 1 | 0.561686 | 0.503 | 0.4795 |
| ☐ | ☑ | Batch{1&2-3} | -0.1552816 | 1 | 3.086383 | 2.772 | 0.0981 |
| ☐ | ☐ | Batch{1-2} | 0 | 1 | 0.129409 | 0.115 | 0.7345 |
| ☐ | ☐ | Shift{morning&evening-afternoon} | 0 | 1 | 0.250019 | 0.223 | 0.6373 |
| ☐ | ☐ | Shift{morning-evening} | 0 | 2 | 0.253524 | 0.112 | 0.8938 |
| ☐ | ☐ | Method{Jones-Smith} | 0 | 1 | 1.893824 | 1.709 | 0.1932 |
| ☑ | ☐ | Temperature | 0 | 1 | 0.200745 | 0.179 | 0.6727 |
| ☐ | ☐ | % Humidity | 0 | 1 | 0.82395 | 0.739 | 0.3916 |
| ☐ | ☐ | pH | 0 | 1 | 0.373313 | 0.334 | 0.5644 |
| ☐ | ☐ | Duration | 0 | 1 | 0.013384 | 0.012 | 0.9132 |
| ☐ | ☐ | Agitation | 0 | 1 | 0.041258 | 0.037 | 0.8481 |
| ☐ | ☐ | Concentration A | 0 | 1 | 1.263582 | 1.136 | 0.2883 |
| ☐ | ☐ | Concentration B | 0 | 1 | 1.263582 | 1.136 | 0.2883 |

Step History

| Step | Parameter | Action | "Sig Prob" | Seq SS | RSquare | Cp | p |
|---|---|---|---|---|---|---|---|
| 1 | pH | Entered | 0.5521 | 0.407202 | 0.0025 | 0.0326 | 2 |
| 2 | Batch{1&2-3} | Entered | 0.0907 | 3.28029 | 0.0226 | -0.79 | 3 |
| 3 | pH | Removed | 0.4672 | 0.601109 | 0.0189 | -2.273 | 2 |
| 4 | Lab tech{Bill&Maria-Sarah&John} | Entered | 0.1002 | 3.048739 | 0.0376 | -2.897 | 3 |

# JMP Partitioning Platform

- Interactive analysis allows you to find significant regression variables using “tree” methods
- Split a branch by the most significant variables or prune the least significant variables
- Split, prune or lock specific branches of the “tree”

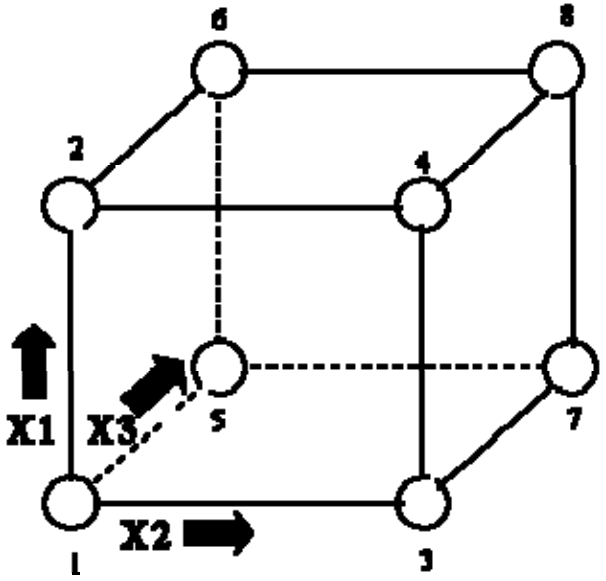


# Design a Screening Experiment

- Use the traditional designs
  - Full factorial and fractional factorial designs
  - Plackett-Burman designs
- Use the JMP DOE menu
  - Screening and Full Factorial menus
  - Custom design menu (*D*-optimal)

# Full Factorial Designs

- Design includes all possible combinations of effects
- Designs become very large as number of factors increases
  - Sample sizes increase more quickly when designs are replicated and center points are added
- Useful for small number of continuous or 2-level factors

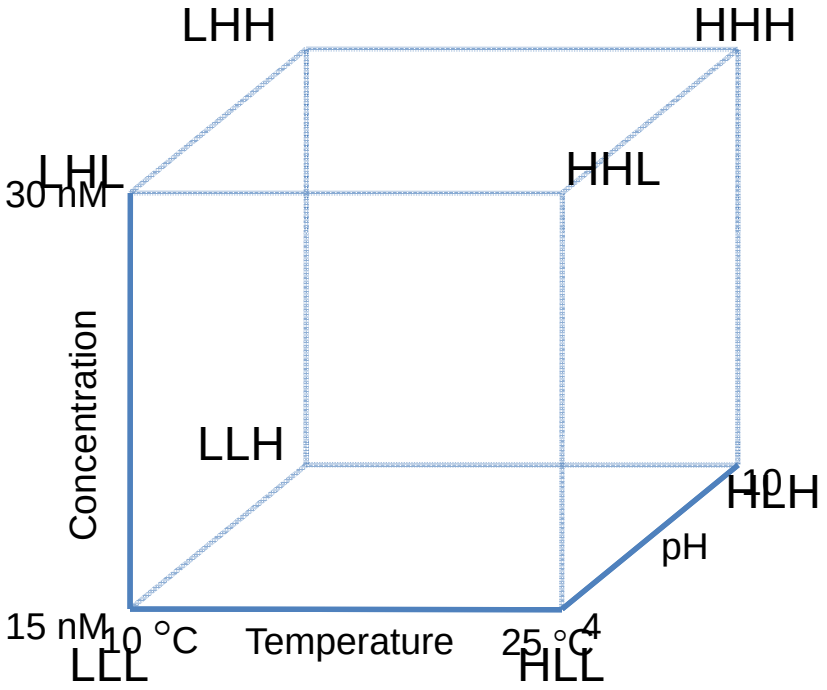



Full Factorial design for three variables with 2 levels each

$2^3$ Factorial design

# Build a $2^k$ Full Factorial Experiment

- A full factorial design experiment is easy to build on paper
  - Small to medium sized factorial designs can be created graphically
  - Larger designs can be created on a spreadsheet or by computer
- Choose high and low values for each factor in the experiment
  - Temperature between 10 and 25 °C
  - Concentration between 15 and 30nM
  - pH between 4 and 10 pH

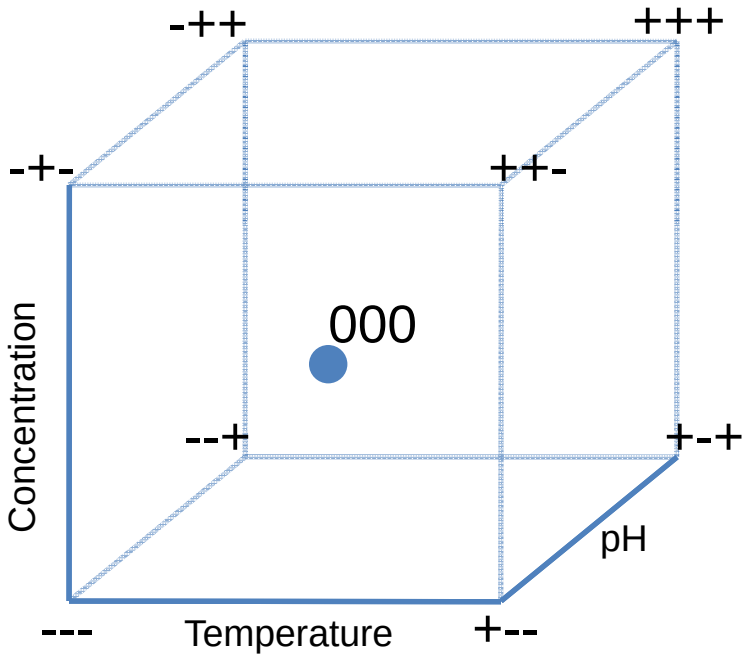



Low points are often marked "-1" or "-"

High points are often marked "1" or "+"

# Center Points

- A center point is sometimes added to a factorial design to estimate curved relationships among factors
  - Need measurements from at least three points to detect a curve
  - Regular factorial designs only use two measurements (high and low)
- A single center point is defined from middle points of each factor
  - Temperature = 17.5
  - Concentration = 22.5
  - pH = 7



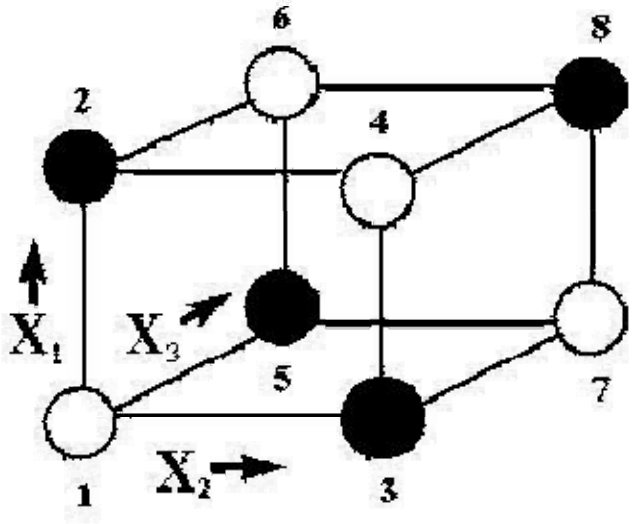

Low points are often marked "-1" or "-"

High points are often marked "1" or "+"

Center points are often marked as "0"

# Fractional Factorial Designs

- Certain combinations of the factors are omitted to create smaller designs
- Omitted design points cause “aliasing” of the interaction and main effects
- Fractional factorial designs exist for different numbers of factors, factor levels and different fractions
  - k = number of factors
  - I = number of factor levels
  - p = fraction of the design




Fractional factorial designs are described using $I^{k-p}$ notation

A $2^{3-1}$ Fractional Factorial design is half the size of a full factorial

# Plackett-Burman Designs

- Very efficient designs for testing a LARGE number of predictor variables with a small number of runs
  - Plackett and Burman. 1946. (Biometrika)
  - Based on Paley construction of Hadamard matrices
- Only allow tests of main effects, not interactions
- Significant interactions are aliased with main effects

# 12 Run Plackett-Burman Design

| | Pattern | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ++++++++++ | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 |
| 2 | -+-+++---+- | -1 | +1 | -1 | +1 | +1 | +1 | -1 | -1 | -1 | +1 | -1 |
| 3 | --+-+++---+ | -1 | -1 | +1 | -1 | +1 | +1 | +1 | -1 | -1 | -1 | +1 |
| 4 | +--+-+++--- | +1 | -1 | -1 | +1 | -1 | +1 | +1 | +1 | -1 | -1 | -1 |
| 5 | -+--+-+++-- | -1 | +1 | -1 | -1 | +1 | -1 | +1 | +1 | +1 | -1 | -1 |
| 6 | --+--+-+++- | -1 | -1 | +1 | -1 | -1 | +1 | -1 | +1 | +1 | +1 | -1 |
| 7 | ---+--+-+++ | -1 | -1 | -1 | +1 | -1 | -1 | +1 | -1 | +1 | +1 | +1 |
| 8 | +---+--+-++ | +1 | -1 | -1 | -1 | +1 | -1 | -1 | +1 | -1 | +1 | +1 |
| 9 | ++---+--+-+ | +1 | +1 | -1 | -1 | -1 | +1 | -1 | -1 | +1 | -1 | +1 |
| 10 | +++---+--+- | +1 | +1 | +1 | -1 | -1 | -1 | +1 | -1 | -1 | +1 | -1 |
| 11 | -+++---+--+ | -1 | +1 | +1 | +1 | -1 | -1 | -1 | +1 | -1 | -1 | +1 |
| 12 | +-+++---+-- | +1 | -1 | +1 | +1 | +1 | -1 | -1 | -1 | +1 | -1 | -1 |

*NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/, date.

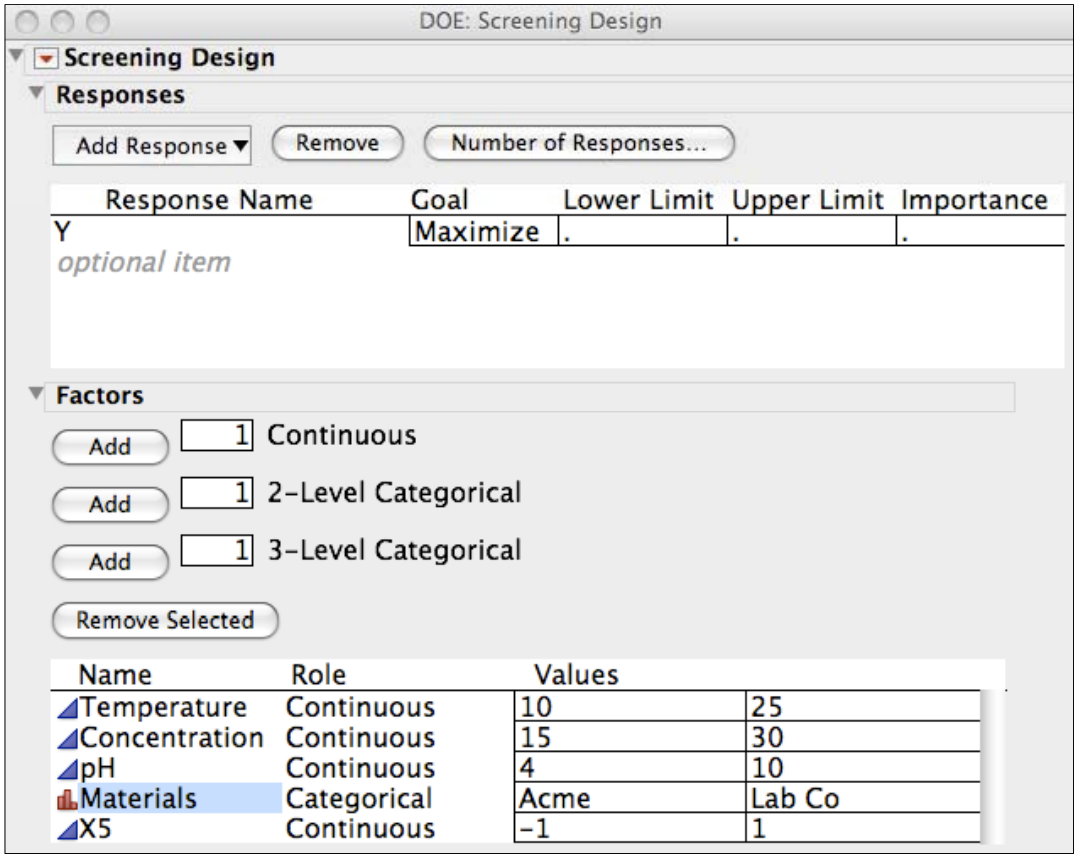
- Each column denotes a unique factor effect
- The -1 and +1 entries represent high and low values and form a Hadamard matrix
- The pattern of -1 and +1 entries represent the Paley construction
- Fewer than 11 factors can be used, if some factors are omitted

# Differences Between Screening Design and Custom Design Menus

- Screening Designs Platform in JMP
  - No blocking factors, covariates or split-plot designs
  - Easy access to effect aliasing information
  - Choose from classical, “named” designs
- Custom Designs Platform in JMP
  - More possible factor types (e.g. blocks, covariates, 4+ level factors, …)
  - Possible to define factor constraints (i.e. restricted factor combinations)
  - Easy to specify required interactions with model dialog
  - Diagnostic plots to explore prediction variance for each design

# JMP Screening Design Menu

- Enter responses and factors using the menu buttons
- Responses can be maximized or minimized or targeted
- Factors can be continuous or categorical with 2-3 levels
  - Enter high and low values
  - Enter factor levels
- Responses and factors can be saved and loaded from the hotspot menu if needed



# JMP Screening Design Menu

Screening Design
7 Factors
Choose a Design

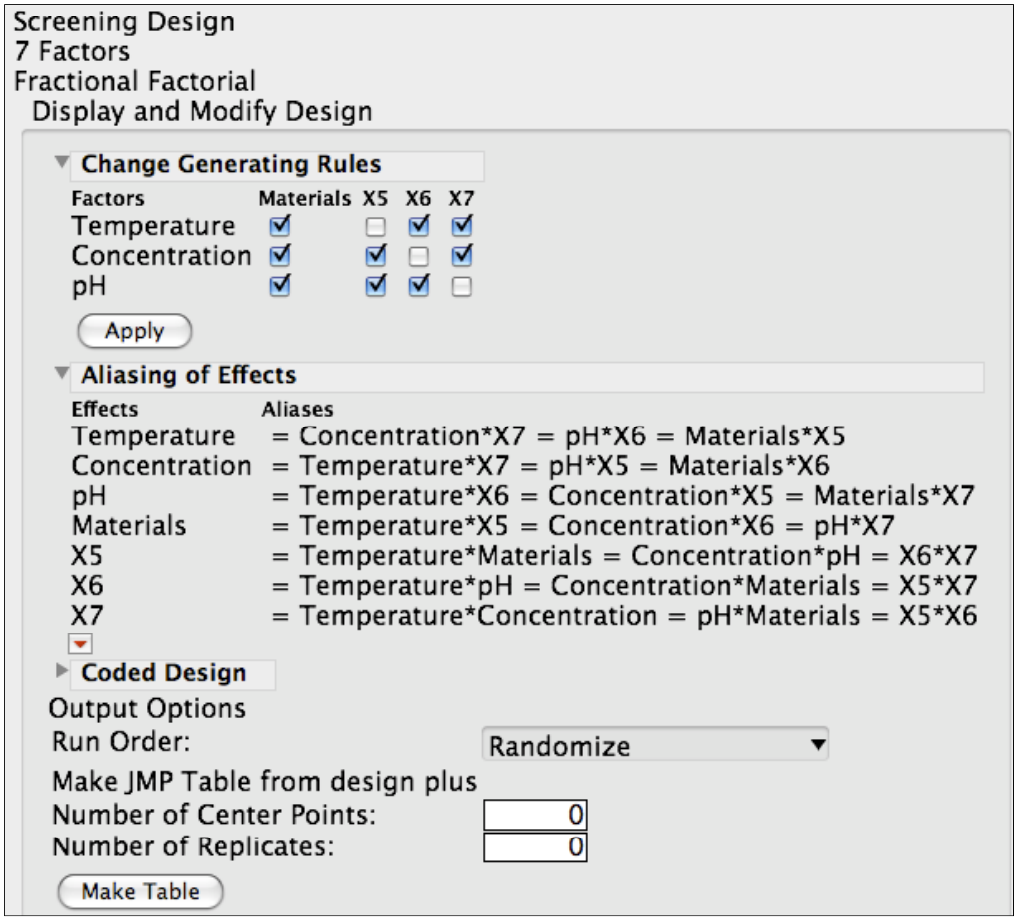
| Number Of Runs | Block Size | Design Type | Resolution – what is estimable |
|---|---|---|---|
| 8 | | Fractional Factorial | 3 – Main Effects Only |
| 12 | | Plackett-Burman | 3 – Main Effects Only |
| 16 | | Fractional Factorial | 4 – Some 2-factor interactions |
| 16 | 8 | Fractional Factorial | 4 – Some 2-factor interactions |
| 16 | 4 | Fractional Factorial | 4 – Some 2-factor interactions |
| 16 | 2 | Fractional Factorial | 4 – Some 2-factor interactions |
| 32 | | Fractional Factorial | 4 – Some 2-factor interactions |

Continue

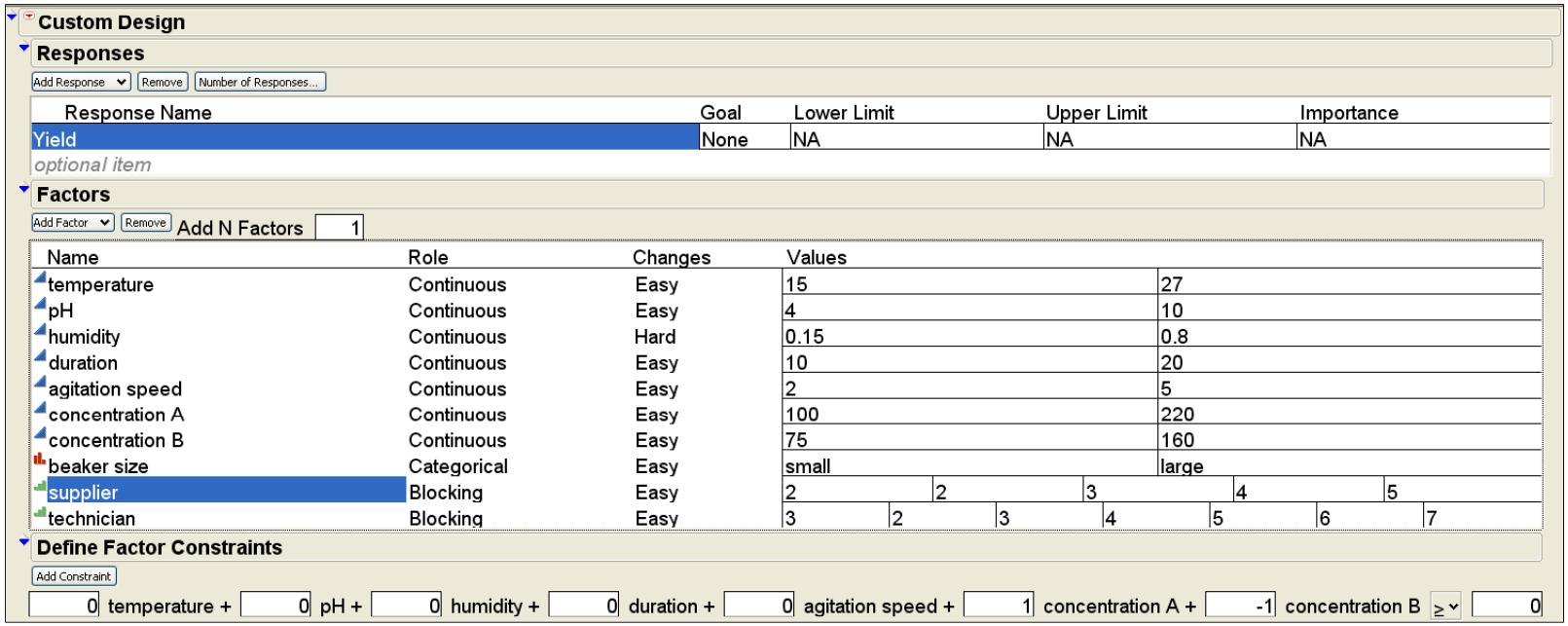
- Select a screening design from the generated list
  - Use number of runs, block size and design preferences to make choice
- Resolution describes how the effects are aliased

# Aliasing in Fractional Designs

- Some design choices will allow you to view the **Aliasing of Effects**
- Interactions and main effects that are aliased cannot be distinguished
- Use **Change Generating Rules** menu to generate alternative alias patterns
- Add center points and replicates then make table



# JMP Custom Design Menu



**Custom Design**

**Responses**

Add Response | Remove | Number of Responses...

| Response Name | Goal | Lower Limit | Upper Limit | Importance |
|---|---|---|---|---|
| Yield | None | NA | NA | NA |

*optional item*

**Factors**

Add Factor | Remove | Add N Factors 1

| Name | Role | Changes | Values | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| temperature | Continuous | Easy | 15 | 27 | | | | | |
| pH | Continuous | Easy | 4 | 10 | | | | | |
| humidity | Continuous | Hard | 0.15 | 0.8 | | | | | |
| duration | Continuous | Easy | 10 | 20 | | | | | |
| agitation speed | Continuous | Easy | 2 | 5 | | | | | |
| concentration A | Continuous | Easy | 100 | 220 | | | | | |
| concentration B | Continuous | Easy | 75 | 160 | | | | | |
| beaker size | Categorical | Easy | small | large | | | | | |
| supplier | Blocking | Easy | 2 | 2 | 3 | 4 | 5 | | |
| technician | Blocking | Easy | 3 | 2 | 3 | 4 | 5 | 6 | 7 |

**Define Factor Constraints**

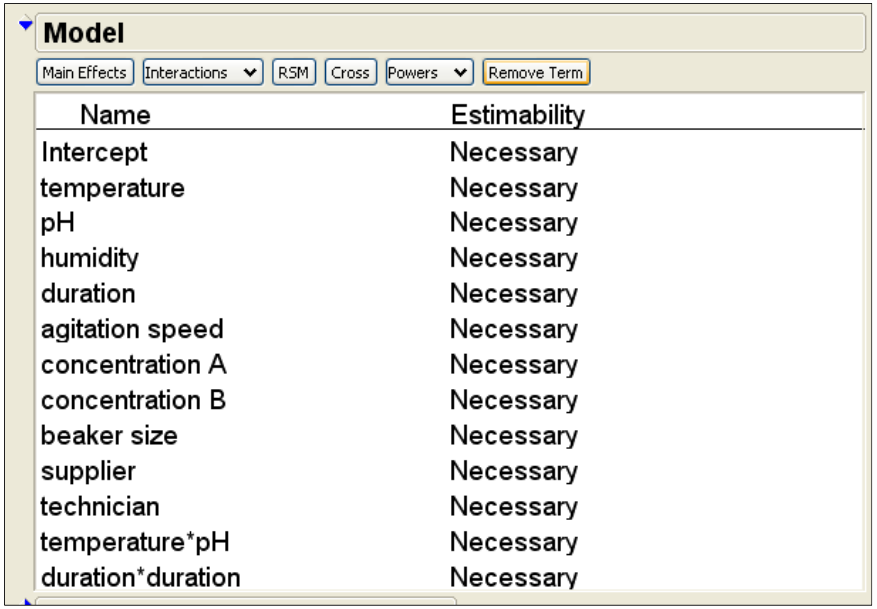
Add Constraint

0 temperature + 0 pH + 0 humidity + 0 duration + 0 agitation speed + 1 concentration A + -1 concentration B ≥ 0

- Choose the D-optimality criteria from the hotspot menu
- Add response(s) and factor variables to the design menus
- Define factor constraints to restrict factor level combinations

# JMP Custom Design Menu

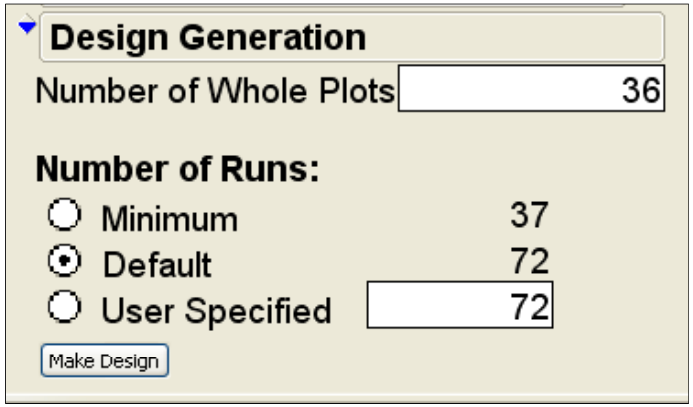
- Specify the effects you need to estimate
- Screening designs focus on main effects
- Add interaction and polynomial effects using menu buttons

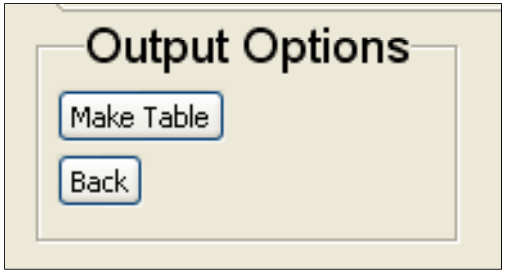



- ***Interactions*** button adds all possible 2-way, 3-way or higher interaction effects
- ***RSM*** = Response Surface Model
- ***Cross*** button for individual interactions
- ***Powers*** button adds all possible second, third or higher degree polynomial effects

# JMP Custom Design Menu

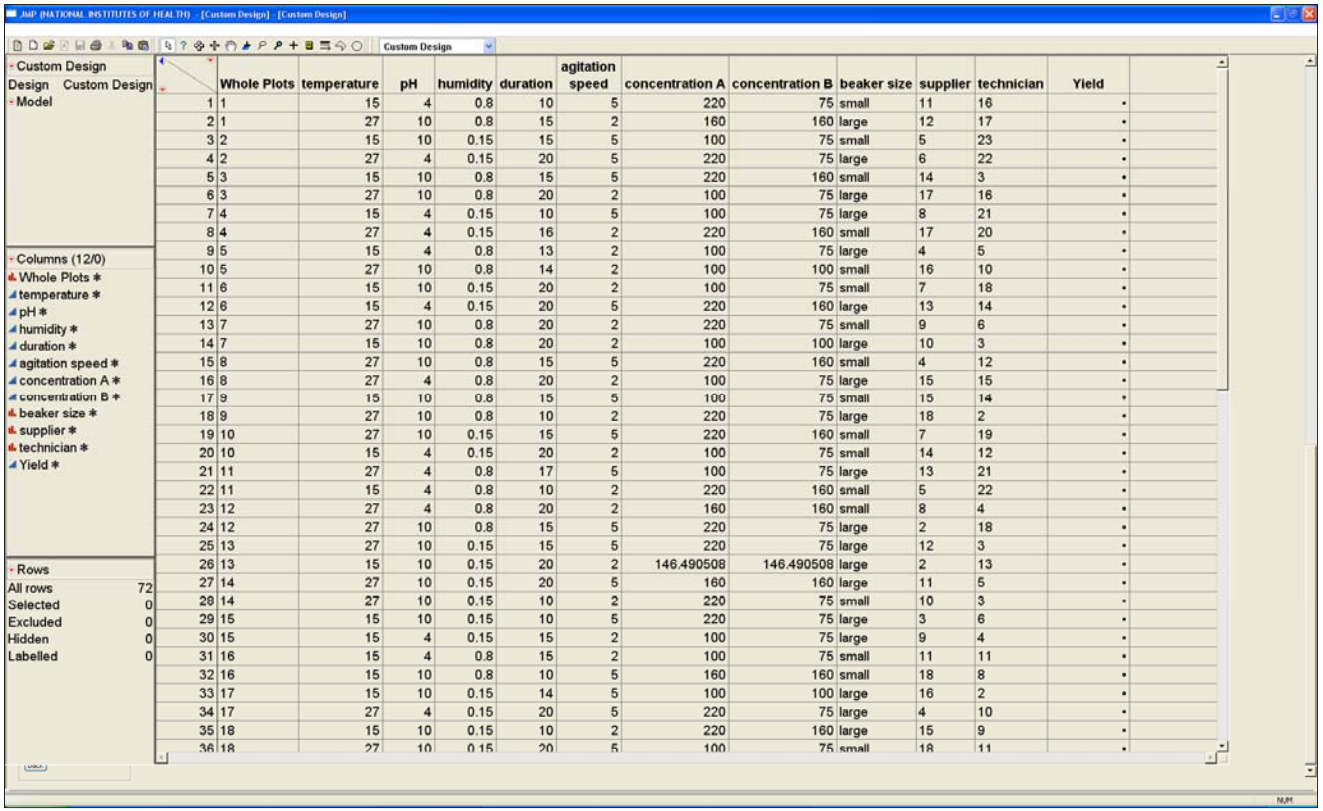
- Choose the number of runs and click ***Make Design***
- JMP will display a design table in the report window
- Click ***Make Table*** to create a JMP data table with your experimental design



# DOE Data Table Output

- JMP completes the data table for you
- Effect variables are randomly assigned to samples (i.e. rows)
- Complete the lab experiments and fill in the response values
- Analyze data from the ***Fit Model*** menu
- Fit model menu will generate effects from your specific design



Custom Design
Design Custom Design
Model

Columns (12/0)
Whole Plots *
temperature *
pH *
humidity *
duration *
agitation speed *
concentration A *
concentration B *
beaker size *
supplier *
technician *
Yield *

Rows
All rows 72
Selected 0
Excluded 0
Hidden 0
Labelled 0

| | Whole Plots | temperature | pH | humidity | duration | agitation speed | concentration A | concentration B | beaker size | supplier | technician | Yield |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 15 | 4 | 0.8 | 10 | 5 | 220 | 75 | small | 11 | 16 | . |
| 2 | 1 | 27 | 10 | 0.8 | 15 | 2 | 160 | 160 | large | 12 | 17 | . |
| 3 | 2 | 15 | 10 | 0.15 | 15 | 5 | 100 | 75 | small | 5 | 23 | . |
| 4 | 2 | 27 | 4 | 0.15 | 20 | 5 | 220 | 75 | large | 6 | 22 | . |
| 5 | 3 | 15 | 10 | 0.8 | 15 | 5 | 220 | 160 | small | 14 | 3 | . |
| 6 | 3 | 27 | 10 | 0.8 | 20 | 2 | 100 | 75 | large | 17 | 16 | . |
| 7 | 4 | 15 | 4 | 0.15 | 10 | 5 | 100 | 75 | large | 8 | 21 | . |
| 8 | 4 | 27 | 4 | 0.15 | 16 | 2 | 220 | 160 | small | 17 | 20 | . |
| 9 | 5 | 15 | 4 | 0.8 | 13 | 2 | 100 | 75 | large | 4 | 5 | . |
| 10 | 5 | 27 | 10 | 0.8 | 14 | 2 | 100 | 100 | small | 16 | 10 | . |
| 11 | 6 | 15 | 10 | 0.15 | 20 | 2 | 100 | 75 | small | 7 | 18 | . |
| 12 | 6 | 15 | 4 | 0.15 | 20 | 5 | 220 | 160 | large | 13 | 14 | . |
| 13 | 7 | 27 | 10 | 0.8 | 20 | 2 | 220 | 75 | small | 9 | 6 | . |
| 14 | 7 | 15 | 10 | 0.8 | 20 | 2 | 100 | 100 | large | 10 | 3 | . |
| 15 | 8 | 27 | 10 | 0.8 | 15 | 5 | 220 | 160 | small | 4 | 12 | . |
| 16 | 8 | 27 | 4 | 0.8 | 20 | 2 | 100 | 75 | large | 15 | 15 | . |
| 17 | 9 | 15 | 10 | 0.8 | 15 | 5 | 100 | 75 | small | 15 | 14 | . |
| 18 | 9 | 27 | 10 | 0.8 | 10 | 2 | 220 | 75 | large | 18 | 2 | . |
| 19 | 10 | 27 | 10 | 0.15 | 15 | 5 | 220 | 160 | small | 7 | 19 | . |
| 20 | 10 | 15 | 4 | 0.15 | 20 | 2 | 100 | 75 | small | 14 | 12 | . |
| 21 | 11 | 27 | 4 | 0.8 | 17 | 5 | 100 | 75 | large | 13 | 21 | . |
| 22 | 11 | 15 | 4 | 0.8 | 10 | 2 | 220 | 160 | small | 5 | 22 | . |
| 23 | 12 | 27 | 4 | 0.8 | 20 | 2 | 160 | 160 | small | 8 | 4 | . |
| 24 | 12 | 27 | 10 | 0.8 | 15 | 5 | 220 | 75 | large | 2 | 18 | . |
| 25 | 13 | 27 | 10 | 0.15 | 15 | 5 | 220 | 75 | large | 12 | 3 | . |
| 26 | 13 | 15 | 10 | 0.15 | 20 | 2 | 146.490508 | 146.490508 | large | 2 | 13 | . |
| 27 | 14 | 27 | 10 | 0.15 | 20 | 5 | 160 | 160 | large | 11 | 5 | . |
| 28 | 14 | 27 | 10 | 0.15 | 10 | 2 | 220 | 75 | small | 10 | 3 | . |
| 29 | 15 | 15 | 10 | 0.15 | 10 | 5 | 220 | 75 | large | 3 | 6 | . |
| 30 | 15 | 15 | 4 | 0.15 | 15 | 2 | 100 | 75 | large | 9 | 4 | . |
| 31 | 16 | 15 | 4 | 0.8 | 15 | 2 | 100 | 75 | small | 11 | 11 | . |
| 32 | 16 | 15 | 10 | 0.8 | 10 | 5 | 160 | 160 | small | 18 | 8 | . |
| 33 | 17 | 15 | 10 | 0.15 | 14 | 5 | 100 | 100 | large | 16 | 2 | . |
| 34 | 17 | 27 | 4 | 0.15 | 20 | 5 | 220 | 75 | large | 4 | 10 | . |
| 35 | 18 | 15 | 10 | 0.15 | 10 | 2 | 220 | 160 | large | 15 | 9 | . |
| 36 | 18 | 27 | 10 | 0.15 | 20 | 5 | 100 | 75 | small | 18 | 11 | . |

# Effect Screening Emphasis

- Choose effect screening emphasis at **Fit Model** menu to generate **Scaled Estimates** report and graph
- Click "hotspot" to generate Normal, Bayes and Pareto plots
- These methods help identify the significant effects and determine the optimal number of variables to include in future experiments

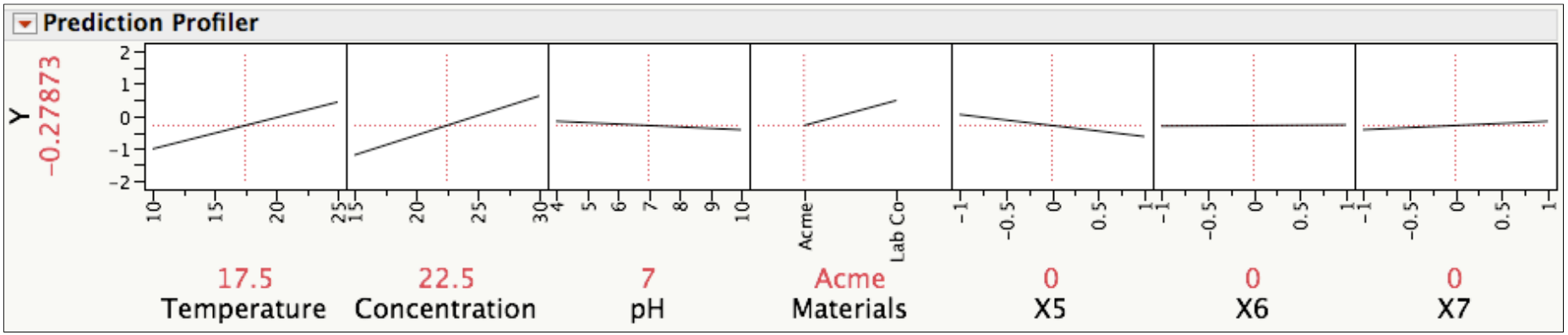
**Scaled Estimates**

Nominal factors expanded to all levels

| Term | Scaled Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | -0.484547 | 0.622543 | -0.78 | 0.4465 |
| temperature(15,27) | 0.1938689 | 0.158422 | 1.22 | 0.2368 |
| pH(4,10) | 0.2741454 | 0.158186 | 1.73 | 0.1002 |
| humidity(0.15,0.8) | -0.443728 | 0.15752 | -2.82 | 0.0114* |
| duration(10,20) | -0.015464 | 0.196963 | -0.08 | 0.9383 |
| agitation speed(2,5) | 0.07513 | 0.157779 | 0.48 | 0.6397 |
| concentration A(0.3,0.7) | -0.524526 | 0.42264 | -1.24 | 0.2305 |
| concentration B(0.5,1) | -0.259809 | 0.50913 | -0.51 | 0.6160 |
| beaker size[small] | -0.110618 | 0.160007 | -0.69 | 0.4982 |
| beaker size[large] | 0.1106184 | 0.160007 | 0.69 | 0.4982 |
| technician[John] | 0.0413061 | 0.158395 | 0.26 | 0.7972 |
| technician[Mary] | -0.041306 | 0.158395 | -0.26 | 0.7972 |
| supplier[ACME] | -0.031924 | 0.157787 | -0.20 | 0.8419 |
| supplier[LabCo] | 0.0319236 | 0.157787 | 0.20 | 0.8419 |
| batch[1] | -0.815441 | 0.474122 | -1.72 | 0.1026 |
| batch[2] | -0.101738 | 0.473209 | -0.21 | 0.8322 |
| batch[3] | -0.950327 | 0.469942 | -2.02 | 0.0583 |
| batch[4] | 0.277162 | 0.471241 | 0.59 | 0.5637 |
| batch[5] | 0.1695993 | 0.474836 | 0.36 | 0.7251 |
| batch[6] | -0.503469 | 0.471241 | -1.07 | 0.2995 |
| batch[7] | 0.9908827 | 0.475461 | 2.08 | 0.0517 |
| batch[8] | 0.3929209 | 0.473209 | 0.83 | 0.4172 |
| batch[9] | 0.7358246 | 0.475461 | 1.55 | 0.1391 |
| batch[10] | -0.195414 | 0.473209 | -0.41 | 0.6845 |
| duration*duration | 0.2939503 | 0.356375 | 0.82 | 0.4203 |
| temperature*humidity | 0.0557553 | 0.158969 | 0.35 | 0.7299 |

Scaled estimates are centered by the mean and scaled by range / 2 for easy comparison of effect sizes

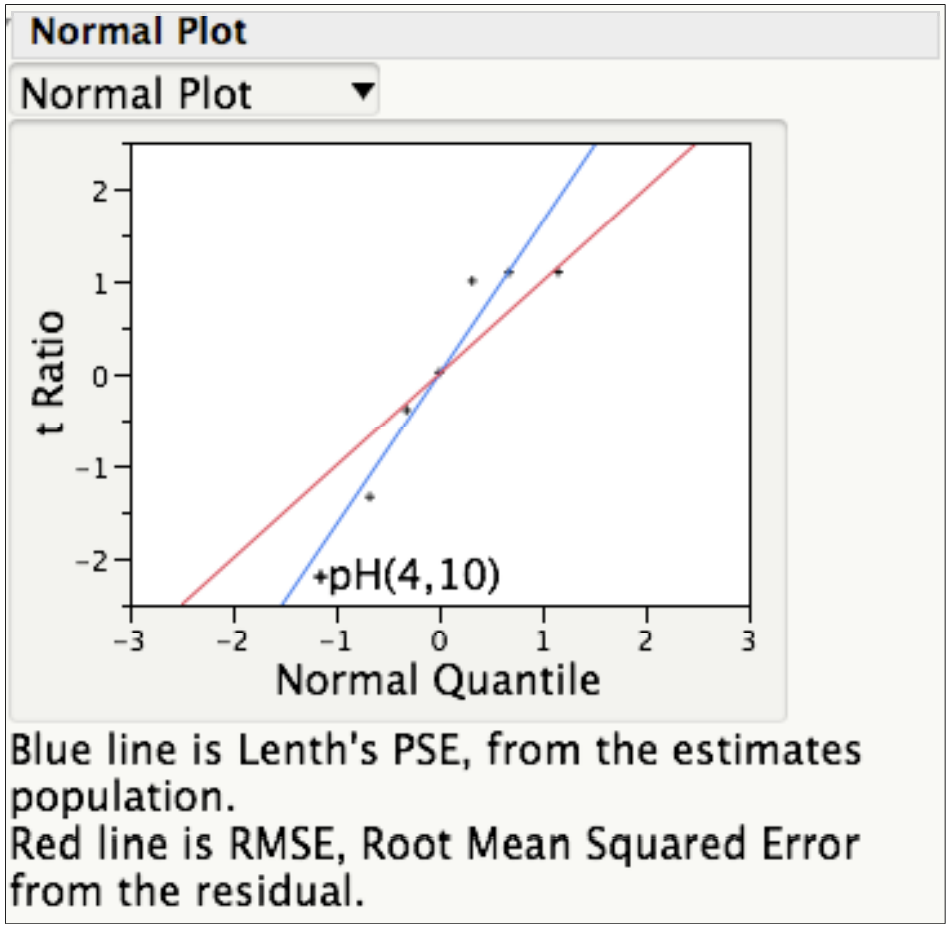
# Prediction Profiler




- Use the prediction profiler to explore the relationships between each factor effect and the response variable
  - Click and drag the red crosshairs to generate predictions
  - Dialogs are interactive to help you explore interaction effects
  - Strong slopes indicate significant factors or large effect sizes

# Normal Plots

- If effects are equivalent to random noise, they should follow a straight line with a slope equal to the variance
  - Red line estimates σ with root mean square error (RMSE)
  - Blue line estimates σ using Lenth's PSE methods for screening design experiments
  - Half normal plot is plotted against absolute normal dist.
- Significant effects deviate from both lines and will be labeled in the normal plot



# Bayes Plots

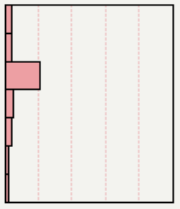
- The Bayes plot estimates the (posterior) probability that an effect is non-zero (significant)
  - Posterior probabilities near one indicate significant effects
  - Posterior probabilities near zero
- Bayes plot menu allows you to specify several initial values
  - **Estimate** is the initial value of each factor effect coefficient
  - **Prior Prob** is the initial estimate of each posterior probability
  - **K Contam** is the magnitude of the significant factor effects

**Bayes Plot**

Response: Y

| K Contam | Std Err Est | DFE |
|---|---|---|
| 10 | 1 | 14 |

| Term | Estimate | Prior | Posterior |
|---|---|---|---|
| Temperature(10,25) | 1.093010 | 0.2000 | 0.0420 |
| Concentration(15,30) | 1.096260 | 0.2000 | 0.0421 |
| pH(4,10) | -2.197491 | 0.2000 | 0.2065 |
| Materials[Acme] | -1.338501 | 0.2000 | 0.0554 |
| X5 | 1.001388 | 0.2000 | 0.0384 |
| X6 | 0.017971 | 0.2000 | 0.0244 |
| X7 | -0.388841 | 0.2000 | 0.0261 |



Posterior Prob that sample is uncontaminated: 0.6325

Enter Prior Prob, Contamination Coef. K, and Click Go

Response: Y

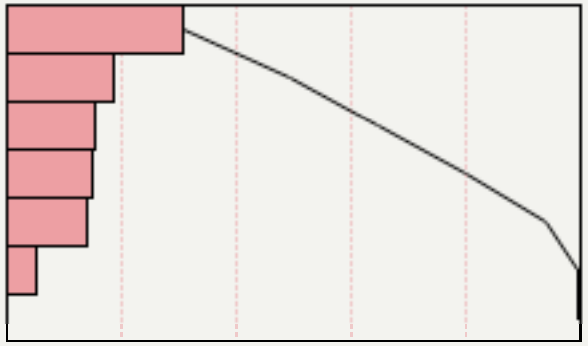
| Term | Estimate | Prior Prob | K Contam | Std Err Scale | DFE |
|---|---|---|---|---|---|
| Temperature(10,25) | 1.09301 | 0.2000 | 10 | 1 | 14 |
| Concentration(15,30) | 1.09626 | 0.2000 | | | |
| pH(4,10) | -2.19749 | 0.2000 | | | |
| Materials[Acme] | -1.33850 | 0.2000 | | | |
| X5 | 1.00139 | 0.2000 | | | |
| X6 | 0.01797 | 0.2000 | | | |
| X7 | -0.38884 | 0.2000 | | | |

Go Help

# Pareto Plots

**Pareto Plot of Estimates**

| Term | t Ratio |
|---|---|
| pH(4,10) | -2.197491 |
| Materials[Acme] | -1.338501 |
| Concentration(15,30) | 1.096260 |
| Temperature(10,25) | 1.093010 |
| X5 | 1.001388 |
| X7 | -0.388841 |
| X6 | 0.017971 |



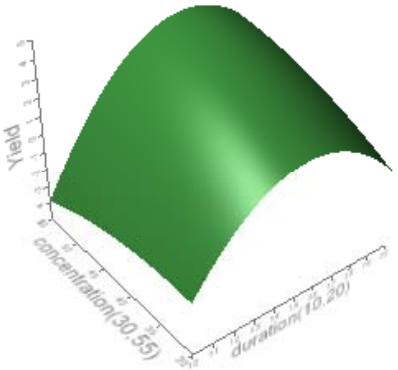
- Pareto plots display absolute values of each orthogonal effect estimates (red bars) and the sum of all effects
- Pareto plots allow you to the size of each effect to the total amount of variance explained by the model

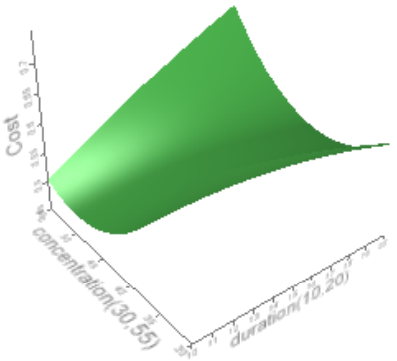
(10 minute break)

# Optimization and Response Surface Design

# Why Response Surface?

- The goal is to maximize, minimize or target a certain value of the response variable(s)
- We must accurately describe the complex multidimensional relationships between the predictors and responses to optimize results
- We use main effects, two-way interactions and 2nd order polynomials to describe these relationships, because they are easiest to interpret and cover most of the relationships



# Example: Fermentation Experiment

- Duration, substrate concentration and temperature are the most significant effects in a fermentation process
- Want to simultaneously maximize yield, minimize costs and target a specific mass of the fermentation product
- Need an efficient experimental design and analytical tools to optimize the fermentation process

# Design Response Surface Experiments

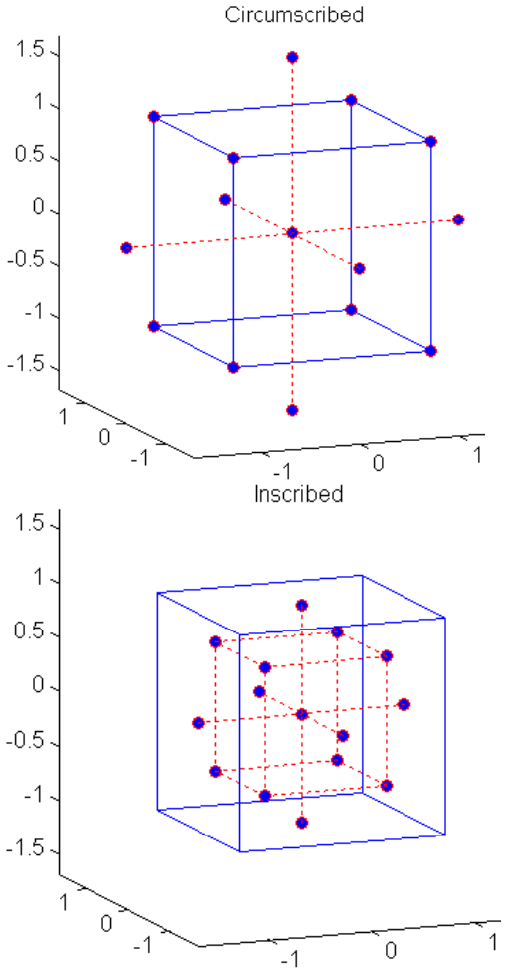
- Use the traditional designs
  - Central Composite Designs
  - Box-Behnken designs

- Use the JMP DOE menu
  - Response Surface menu
  - Custom design menu (*I*-optimal)

# Central Composite Designs

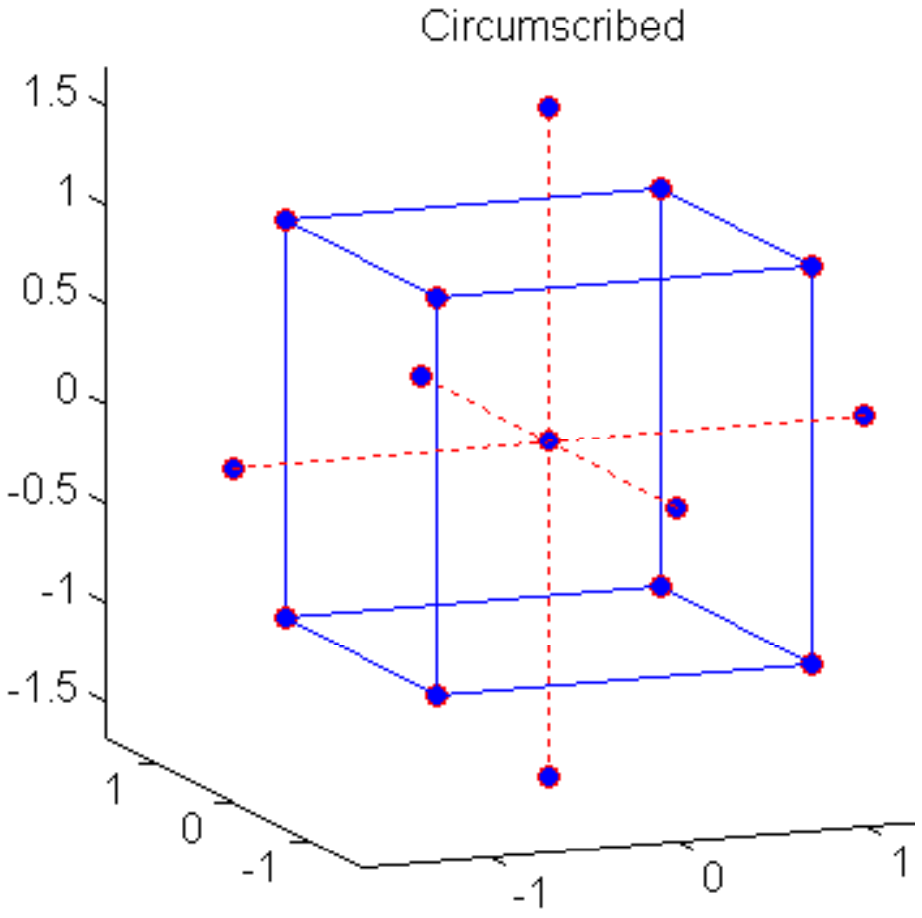
- Central composite designs are full factorial designs augmented with center points and axial 'star' points to evaluate polynomial effects

- Three types of CC designs are distinguished by their axial points
  - Circumscribed Central Composite (CCC)
  - Inscribed Central Composite (CCI)
  - Face-centered Central Composite (CCF)



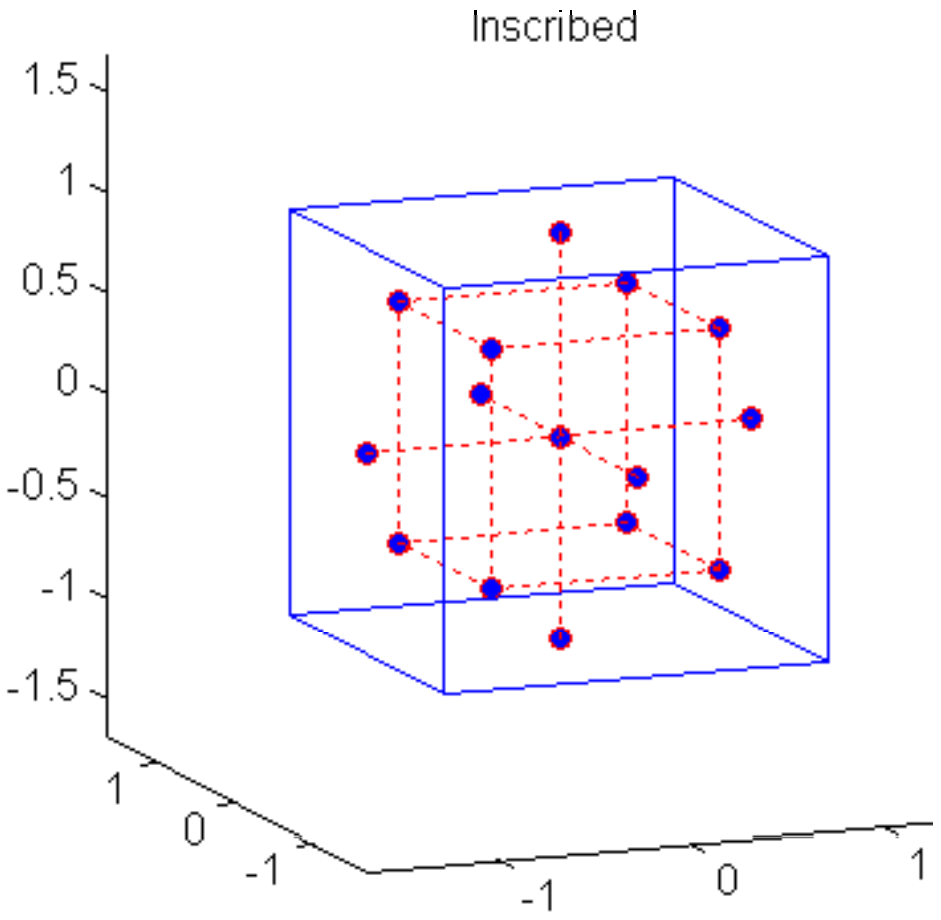
# Circumscribed (CCC) Designs

- Axial points are outside the design space of a regular full factorial design experiment
- Advantages:
  - Highest quality predictions
  - Easy to create CCC design from a full factorial experiment
- Disadvantages:
  - Axial points may include some unreasonable factor values
  - E.g. if full factorial design uses pH from 4 to 10, then axial points may be pH = 2 and 12



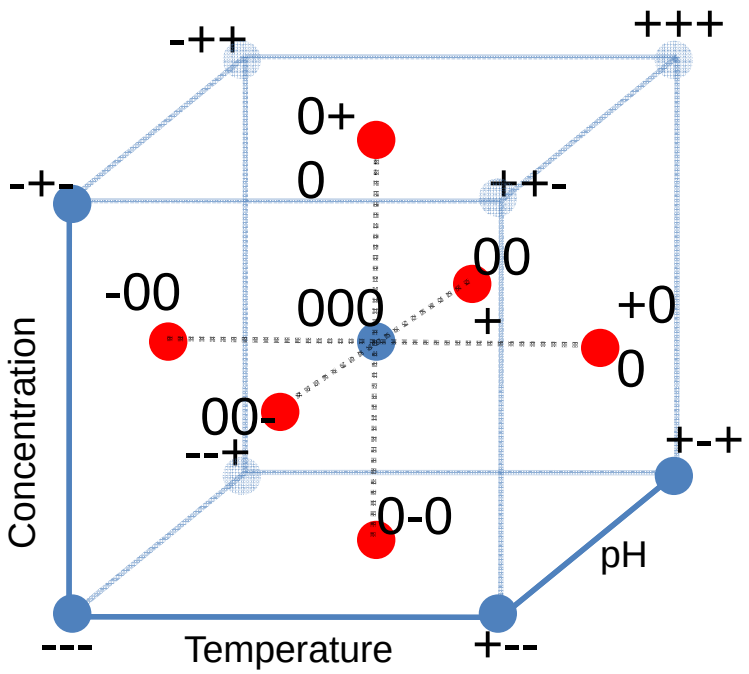
# Inscribed (CCI) Designs

- Smaller range of values used in full factorial design for more reasonable axial point values
- Advantages:
  - Reasonable axial point values
  - Very high quality predictions
- Disadvantages:
  - Cannot create CCI design from an existing full factorial design
  - Slightly lower quality predictions than the CCC design due to the location of the design points



# Face-Centered (CCF) Designs

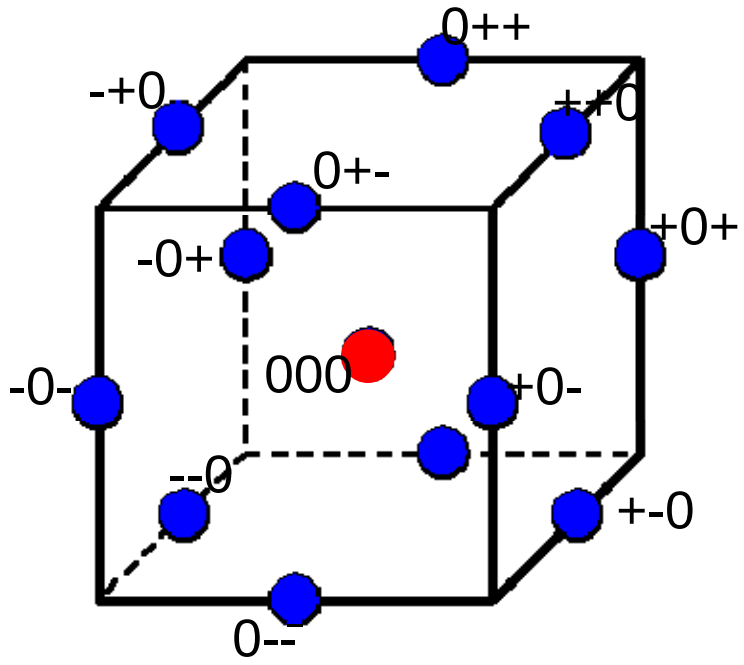
- Axial design points located on faces of the full factorial design "cube"
- Advantages:
  - Reasonable axial point values
  - Mostly high quality predictions
  - Easy to create from full factorial design
- Disadvantages:
  - Poor quality prediction of all pure quadratic effects (e.g. Temperature$^2$)
  - Only 3 levels per factor vs. 5 levels per factor for CCC and CCI designs




Note: You can use the familiar "-", "0" and "+" notation from full factorial designs to describe CCF designs

# Box-Behnken Designs

- Full factorial design is not used, but samples are assigned to the “edges” of a full factorial design space
- Advantages:
  - Smaller samples sizes than CC designs
  - No unreasonable design points
- Disadvantages:
  - Poor predictions in the “corners” of the design space (i.e. extreme combinations)
  - Cannot create a Box-Benkin (BB) design from an existing full factorial design
  - May not be a “rotatable” design




Compare Box-Behnken designs to factorial and central composite designs

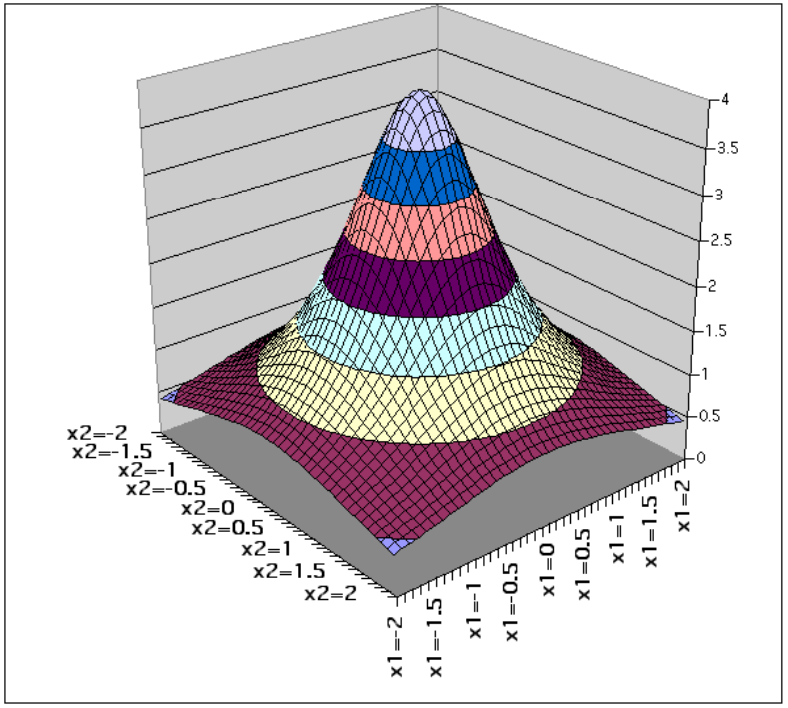
Only 3 levels per factor

# Rotatable Designs

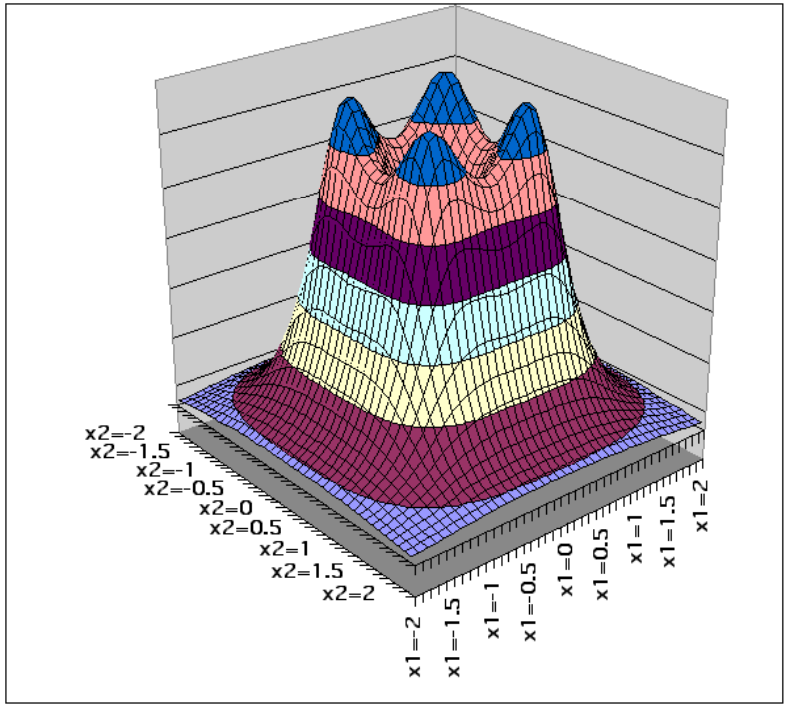
- A designed experiment is ***rotatable*** if its prediction variances are a function of the distance from its center point, but not the direction away from center

- If a design is **NOT** rotatable, it will have regions of increased or decreased precision that may create biased predictions and misleading optimizations

# Rotatable Designs

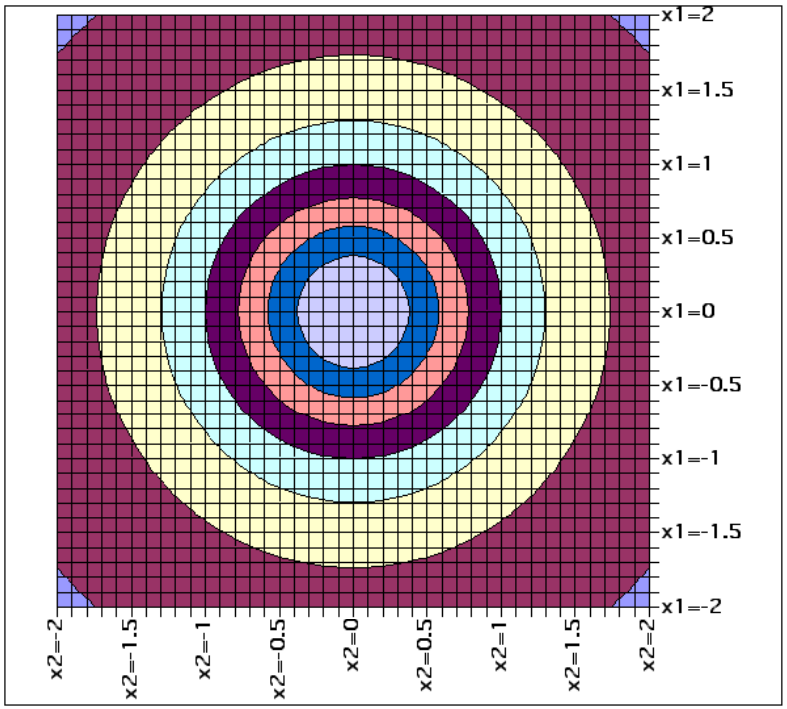



- Rotatable $2^2$ design

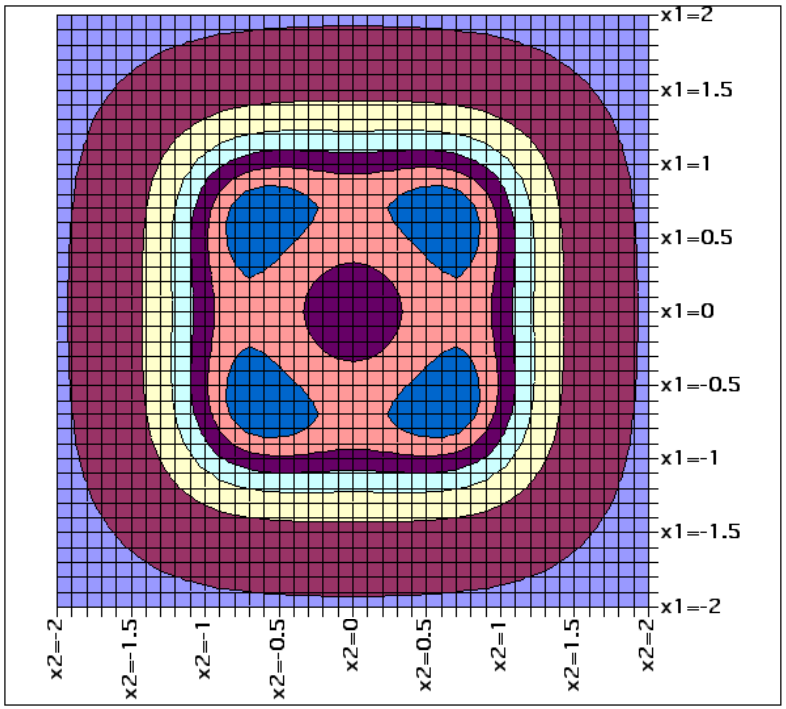



- Non-rotatable $3^2$ design

# Rotatable Designs




- Rotatable $2^2$ design




- Non-rotatable $3^2$ design

# Orthogonal Designs

- An experimental design is orthogonal if all factor main effects and interactions are independent of each other
- Screening designs are usually not orthogonal, because their effects are typically aliased
- Want to use orthogonal designs to produce accurate estimates of coefficients and more precise optimizations in RSM analysis

Recall non-orthogonal effects are often created through poor design choices (e.g. baby gender and blanket color)

We should also be concerned about collinear predictors (e.g. two separate factors for mass and waistline)

# Choosing a Known Design

| Design | Pros | Cons |
|---|---|---|
| CCC | High quality predictions over the entire design space | Large sample sizes and axial points may include unreasonable values |
| CCI | High quality predictions over a slightly smaller design space. Reasonable axial point values. | Large sample sizes and slightly lower quality predictions than CCC |
| CCF | Relatively high quality predictions. Reasonable axial point values. | Large sample sizes and poor prediction of pure quadratic effects |
| BB | Smallest sample sizes, while using reasonable axial point values. | Poor predictions in “corners” of the design space. May not be rotatable. |

- Choosing a known design may require some compromise
  - Modifying a full factorial screening experiment, sample size, …
- JMP Custom Designs may provide better solutions

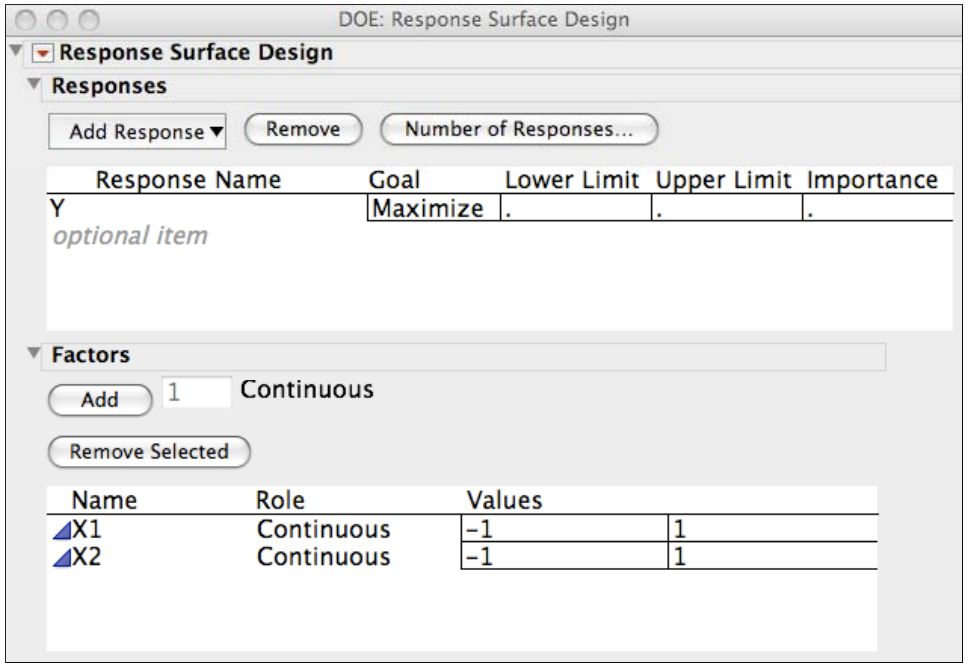
# Design Response Surface Experiments

- Use the traditional designs
  - Central Composite Designs
  - Box-Behnken designs

- Use the JMP DOE menu
  - Response Surface menu
  - Custom design menu (*I*-optimal)

# JMP Response Surface Menu

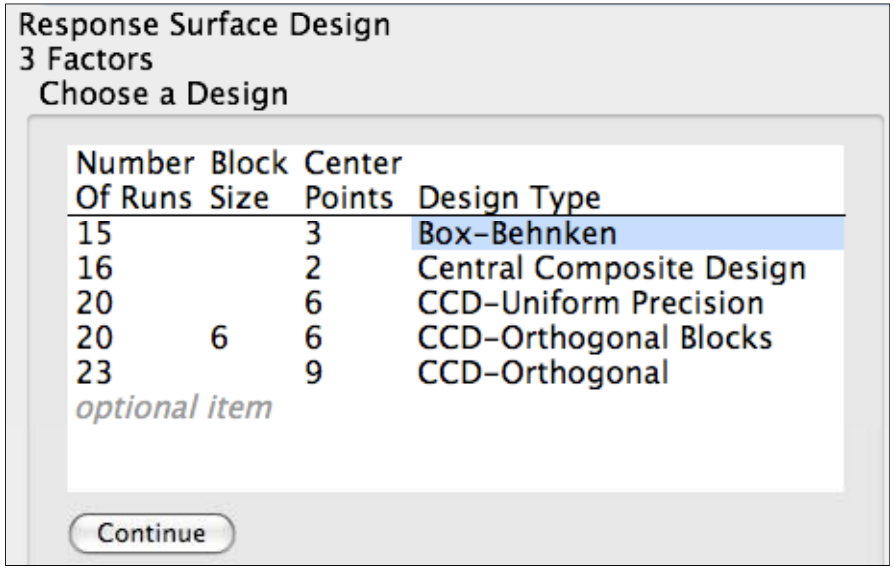
- Enter responses and continuous factors using the menu buttons
- Responses can be maximized, minimized or targeted
  - Use Lower Limit and Upper Limit fields to set optimization boundaries
  - Optimization goals and boundaries are used when you analyze your data
- Responses and factors can be saved and loaded from the hotspot menu if necessary

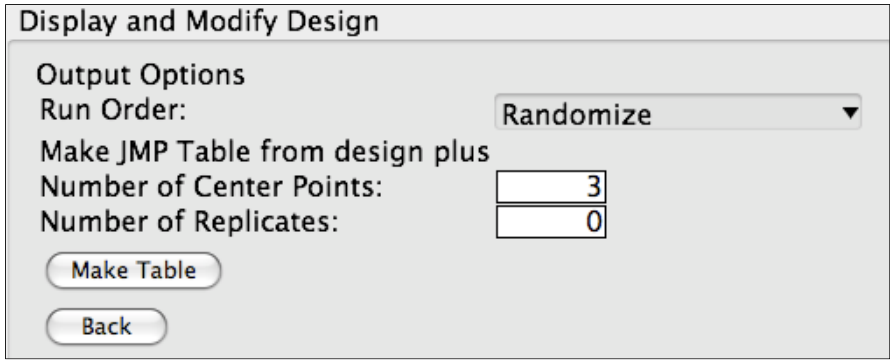



Notice there are no options to add any covariates, categorical or blocking factors

# JMP Response Surface Menu

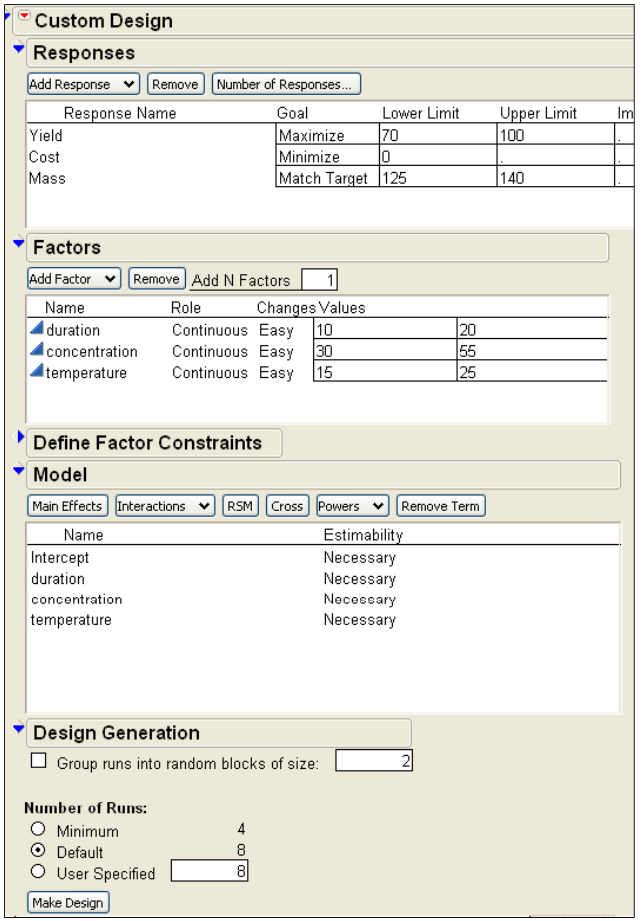
- Select a response surface design from the list
- Choose the design based on number of runs or blocking
- Understand the advantages and limitations of the designs
- Add replicates and/or center points, then make a table



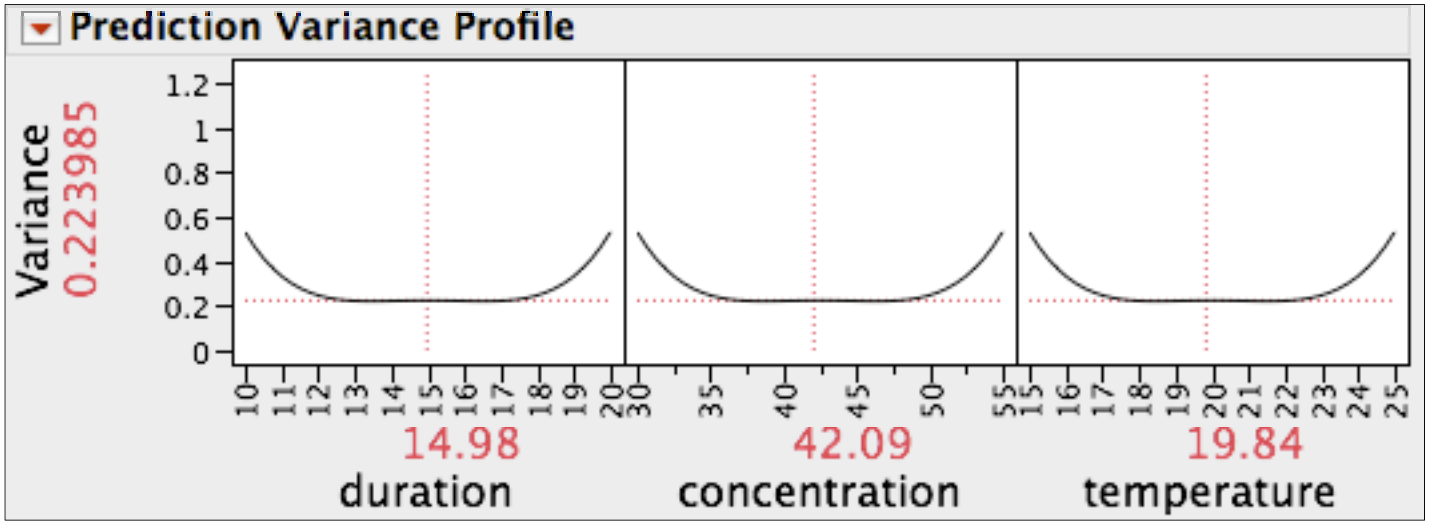



# JMP Custom Design Menu

- Add the responses, factors, constraints and model effects just as before
- Use RSM button to specify a response surface model design and analysis
- Generate the data table as before
- Complete the experiments, enter the response data into the table and click Fit Model to start the optimization



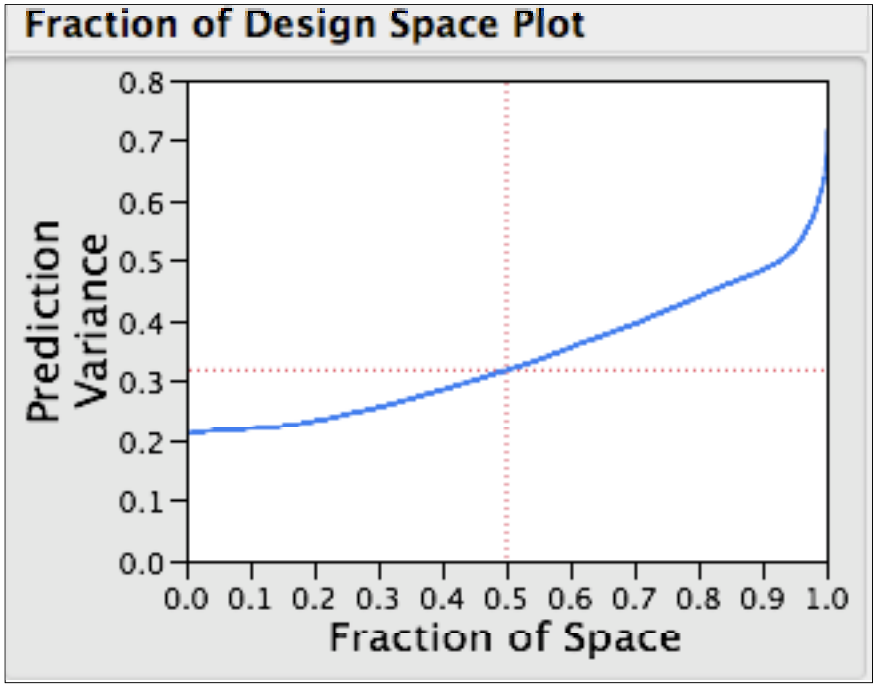
# Design Variance




- The **Prediction Variance Profile** tool allows you to explore how the variance will change over different predictor values
- Prediction variance is reported as a fraction of the error variance, where *Variance = 1* implies *prediction variance = error variance*

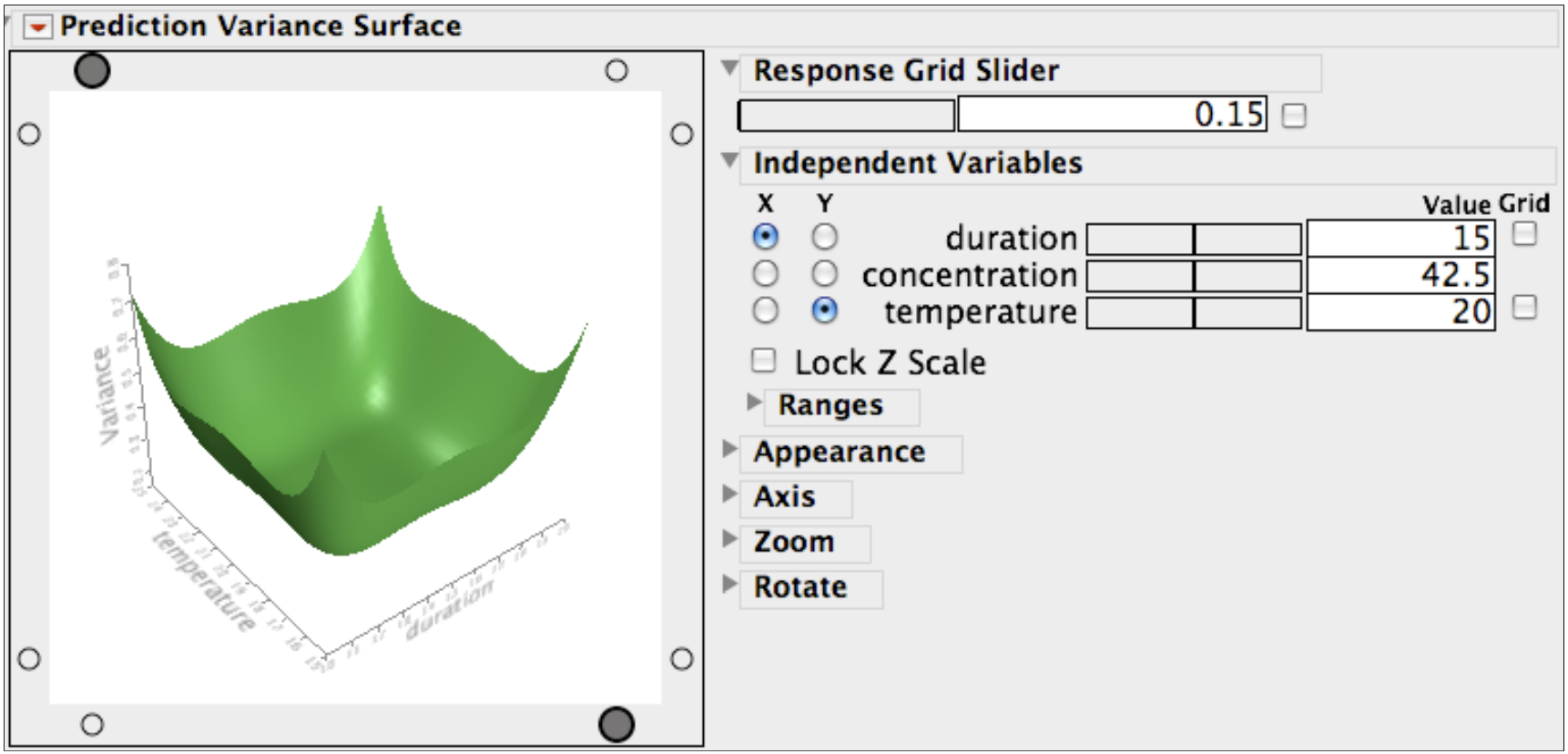
# Fraction of Design Space Plot

- Fraction of design space plot displays the proportion of runs with a specific variance
  - E.g. half of all runs will have a prediction variance of 0.3
  - Want majority of runs to have small variances to ensure quality predictions in RSM analysis
- Use fraction of design space plot to assess the quality of an experimental design quickly




Compare *Fraction of Design Space* plot to *Prediction Variance Profile* plot to identify ***where*** the high variance points are located in the design

# Prediction Variance Surface




- Explore prediction variance in 3 dimensions

# Relative Variance of Coefficients

- Differences in samples size among groups can create differences in variance and power among effects
  - E.g. Generally main effects will have more power than interactions or quadratics

- Variance of Coefficients is most useful for custom designs with unequal runs among factors and levels

**Relative Variance of Coefficients**

| | |
|---|---|
| Significance Level | 0.050 |
| Signal to Noise Ratio | 1.000 |

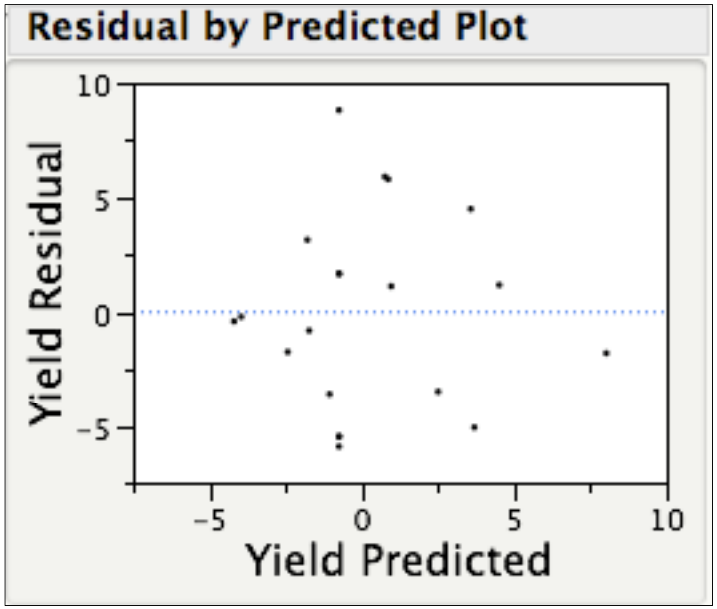
| Effect | Variance | Power |
|---|---|---|
| Intercept | 0.224 | 0.427 |
| duration | 0.100 | 0.750 |
| concentration | 0.100 | 0.750 |
| temperature | 0.100 | 0.750 |
| duration*duration | 0.379 | 0.278 |
| duration*concentration | 0.125 | 0.657 |
| concentration*concentration | 0.379 | 0.278 |
| duration*temperature | 0.125 | 0.657 |
| concentration*temperature | 0.125 | 0.657 |
| temperature*temperature | 0.379 | 0.278 |

Increase Signal to Noise Ratio field to increase power if you anticipate highly significant effects when data is collected

Adjust Significance Level as needed

# Analyzing a RSM Experiment

- Collect data, fill in the JMP data table and analyze with *Fit Model*
- Check usual model assumptions
  - Independent and identically distributed (i.i.d.) normal random errors
  - Look for high influence points
- Improve model fit if necessary
  - Transform or remove predictors
  - Check Lack of Fit test results to see if model is missing important effects

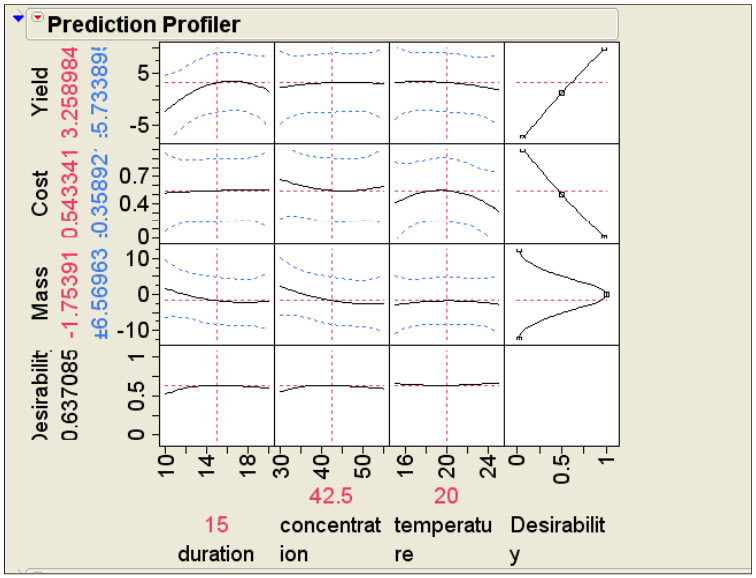



Check residual plot for curved trends or non-constant variance

Check leverage plots to look for high influence points or outliers

# Prediction Profiler

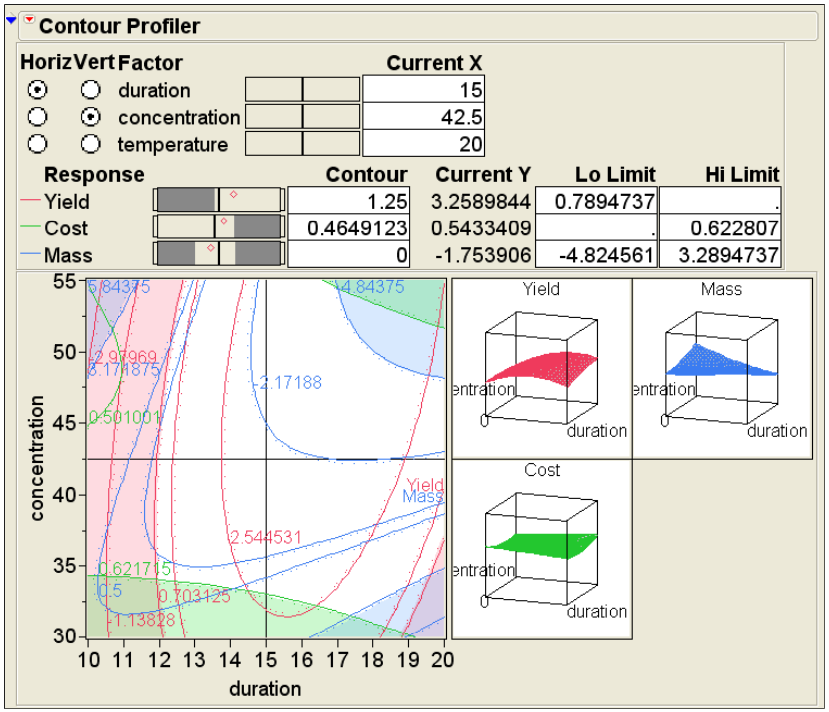
- Add profiler from Least Squares or response hotspot menu
- Add desirability functions from Prediction Profiler hotspot menu
- Desirability plots describe response objectives (e.g. maximize, minimize, match a target value, …)
- Click Maximize Desirability from hotspot menu to find the optimal levels of your predictor variables




- Interactive results from prediction profiler show optimal predictor values and CI's for responses
- Click predictor levels to explore the response surface

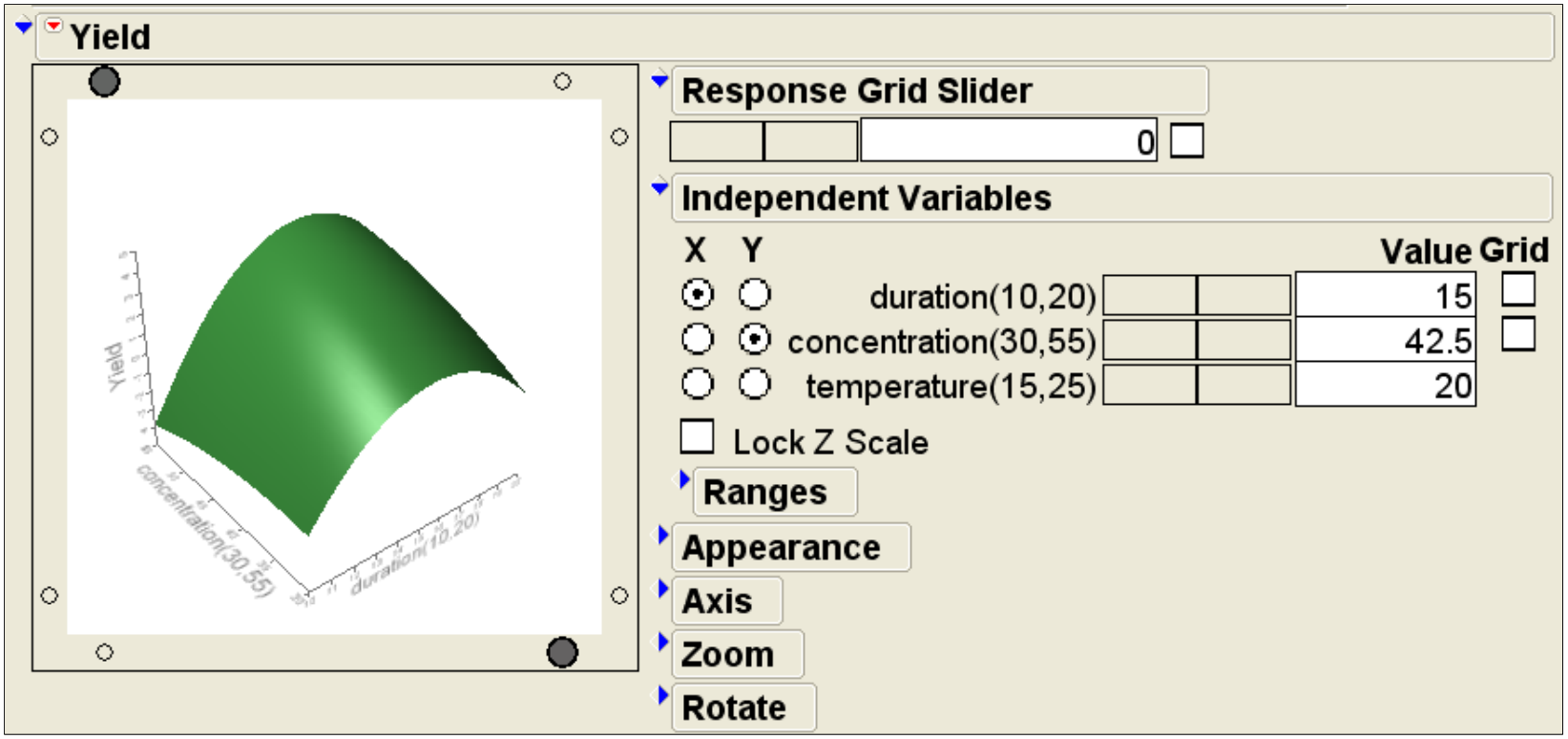
# Contour Profiler

- Add contour profiler from LS or response hotspot menu
- Add contour grids for responses from contour profiler hotspot
- Add response shading from the response variable dialogues
- The white space describes optimized predictor levels




- Prediction profiler is best for point estimates of the optimal predictor levels, while the contour profiler provides a “neighborhood” of optimal predictor variable values

# Surface Profiler




- Another interactive profiler to find optimal predictor variable levels from the RSM

# Literature Cited and Resources


- Plackett and Burman. 1946. The design of optimum multifactor experiments. **Biometrika**. 33(4):305-325

- Paley. 1933. On orthogonal matrices. **J. Math. Phys**. 12:311-320

- *NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/, 2-26-2009