Quality by Design and Biologics Process Development

Mike Fino

MiraCosta College

Western Hub Director, NBC2

Today, in three parts

1. Process development and quality by design (QbD)
2. ANOVA and other statistics we never *really* learned
3. Introduction to design of experiments

Process Development and Quality by Design (QbD)

Section One

Stages of development for a new product

Research

- Discovery
- Preclinical studies

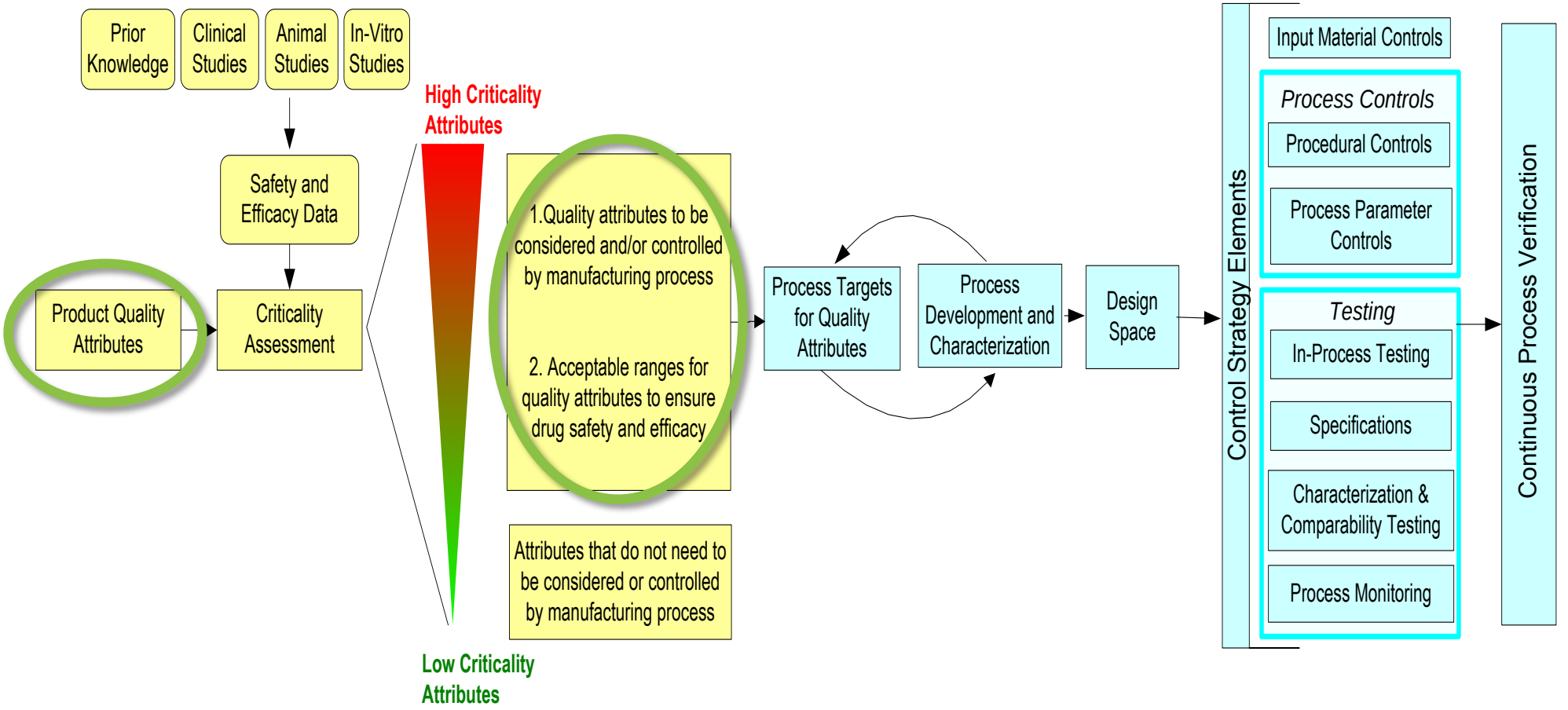
Development

- Clinical studies
- Scale-up

Production

- Quality
- Compliance

Linking Product and Process Understanding



From A-Mab study

Product Quality Attributes

- Identity
- Physicochemical properties
- Quantity
- Potency
- Product-related impurities
- Process-related impurities
- Safety

Product Quality Attributes

- Identity
- Physicochemical properties
- Quantity
- Potency
- Product-related impurities
- Process-related impurities
- Safety

Identity

Strength

Purity

Fundamental Quality Attributes: Monoclonal antibody

- Process-related impurities
  - Host cell proteins
  - DNA
  - Small Molecules
  - Leached Protein A
- Product-related impurities
  - Degradation products
  - Molecular variants with properties different than expected
  - Truncated forms, aggregates
- Safety
  - Microbial load
  - Sterility
  - Endotoxin
  - Mycoplasma and adventitious virus
  - Turbidity
- Quantity
  - Protein content/amount
  - Yield
- Potency
  - Animal, cell, or biochemical assay
- Physicochemical properties
  - Primary structure
  - Higher order structure
  - Molecular weight/size
  - Isoform/charge pattern
- Identity
  - Specific

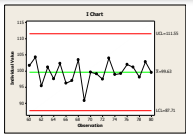
Terminology

- Quality Attributes
  - A physical, chemical, or microbiological property or characteristic of a material that directly or indirectly impacts quality
- Critical Quality Attributes (CQAs)
  - A quality attribute that must be controlled within predefined limits to ensure that the product meets its intended safety, efficacy, stability and performance
  - These are product specific, based on prior knowledge, nonclinical/clinical experience, risk analysis, etc.

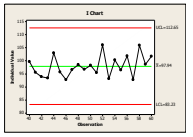
Developing Process Understanding

Process Parameters

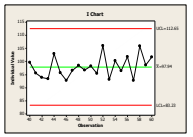
INPUTS (X)



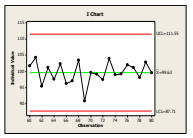
Man



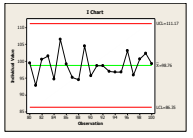
Machine



Methods



Materials



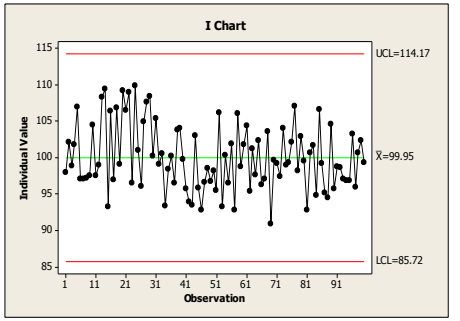
Environment

$y = f(x)$

Quality Attributes

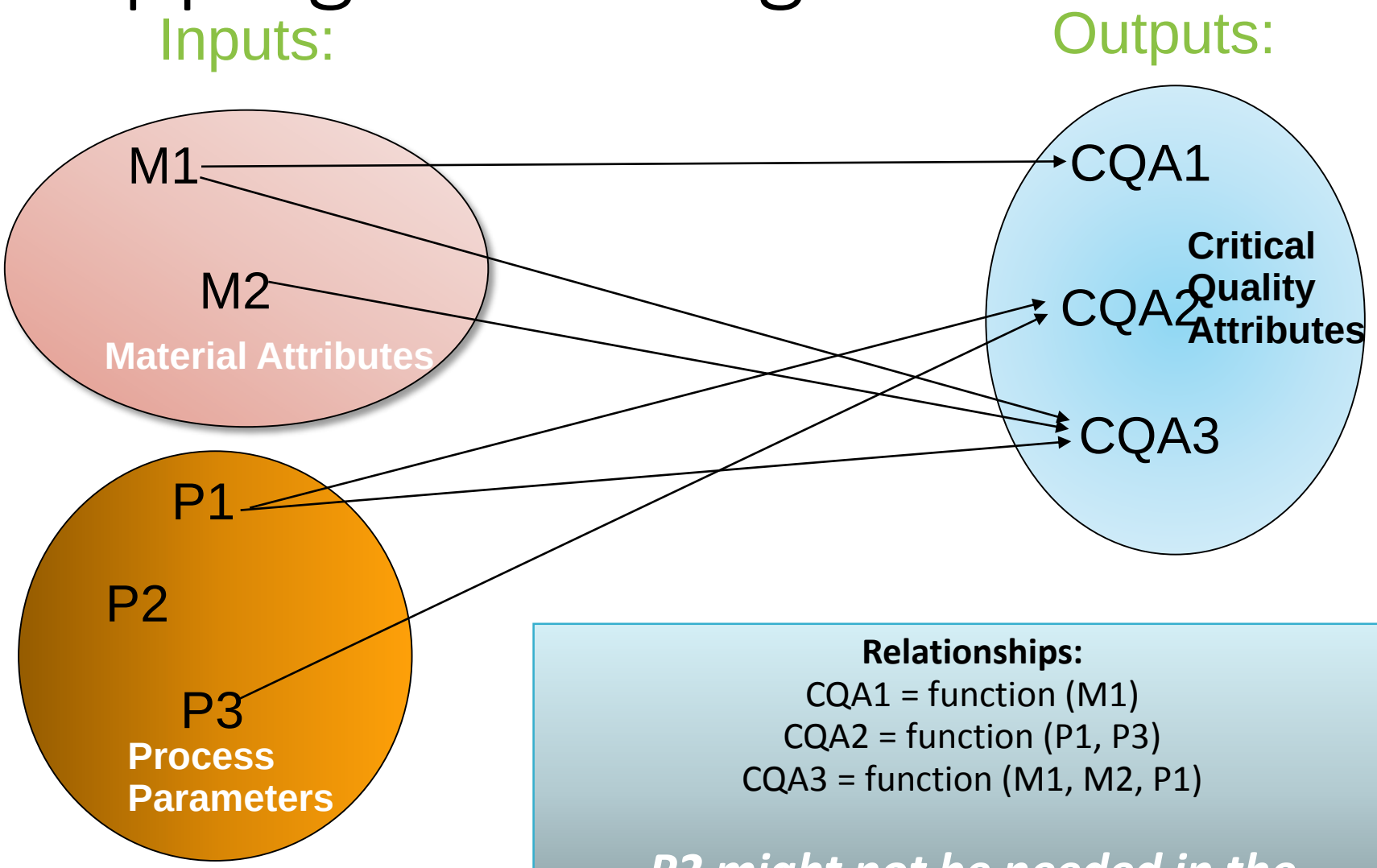
Inputs to the process control variability of the Output

y



OUTPUT

Mapping the Linkage



Relationships:

CQA1 = function (M1)

CQA2 = function (P1, P3)

CQA3 = function (M1, M2, P1)

P2 might not be needed in the establishment of design space

ANOVA and other statistics we never *really* learned

Section Two

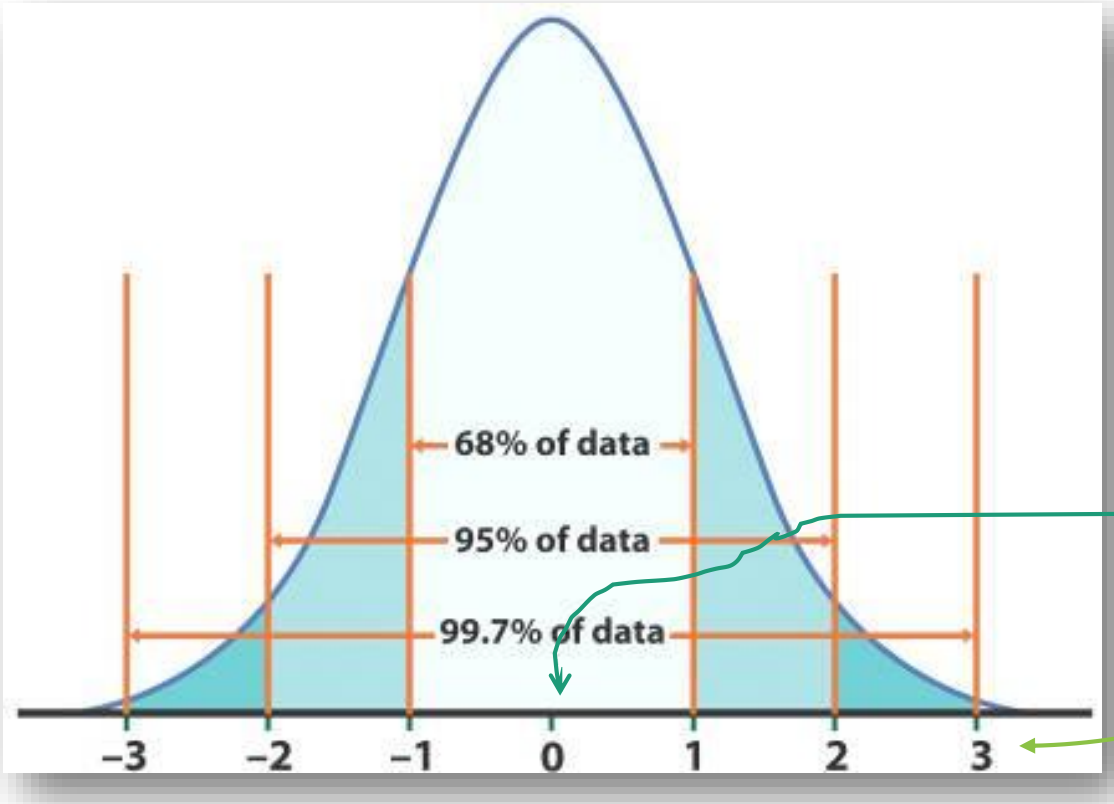
Extending Intro Statistics

- Courses often end with analysis of variance – ANOVA
- ANOVA is all that is needed to understand industrial design of experiments

- Who's comfortable with their knowledge of ANOVA?
  - What can it be used for?
  - What information does it give us?

The Standard Normal

$$z = \frac{(\text{data point} - \text{mean})}{\text{standard deviation}} = \frac{(x_i - \bar{x})}{s}$$



Allows us to work with null model centered on zero

Allows us to see how many standard deviations our observation is from the mean

General form of a test statistic

- There are many different types of test statistics out there and many have the same general form
  - z-score, t-statistic and F-statistic
- General form is a ratio of the difference on top divided by the variability on the bottom

$$\text{test statistic} = \frac{difference}{variability}$$

Standardized Distributions

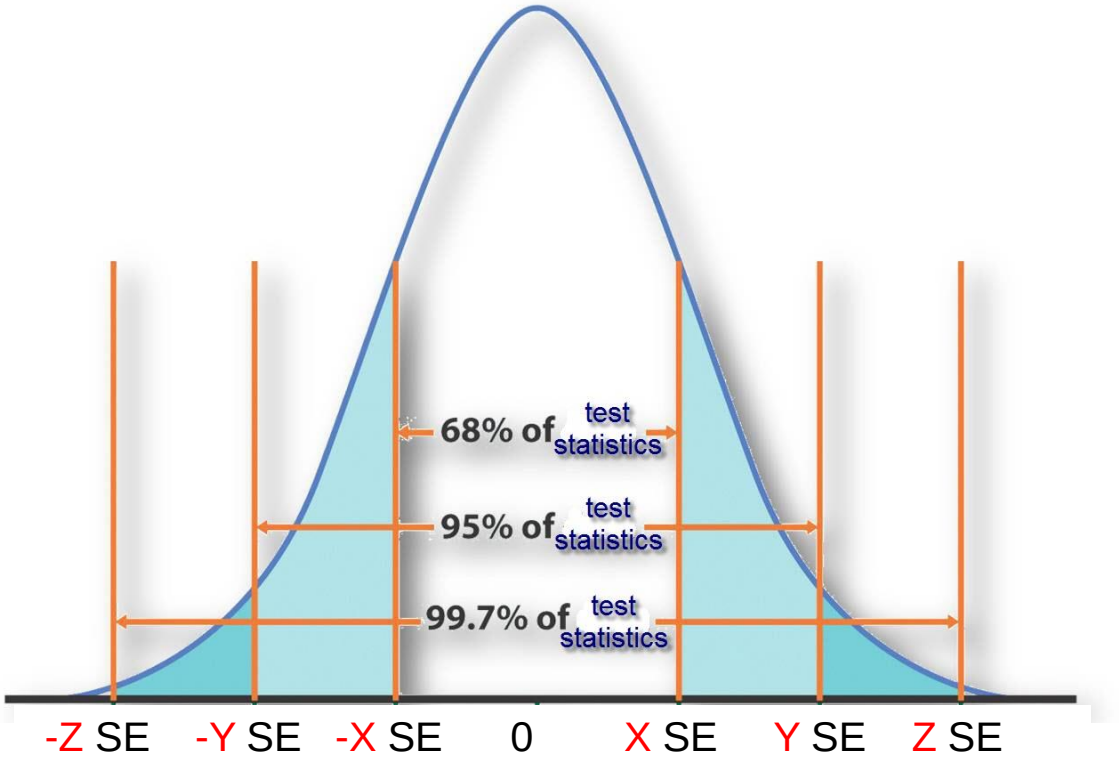
- Standard Normal
  - We use this for individual data (via a z-score)
  - A quick way to see if a data point is unusual or not
- t-distributions
  - We use this for sample means (via a t-statistic)
  - Used in methods to determine if a sample mean is different from the null (one-sample t-test) or if two groups are difference (two-sample t-test)
- F-distributions
  - We use this for sample means (via a F-statistic)
  - Used in methods to determine if two or more sample means are different (ANOVA)

Our Approach to Hypothesis Testing

- Model what the data would like, if the null were true
- Compare our ~~actual results~~ *results wrapped up in a test statistic* to the null
- Ask whether our data would be expected or unexpected in the model
  - Expected data supports the null (e.g. p-value greater than 5%)
  - Unexpected data rejects the null (e.g. p-value less than 5%)

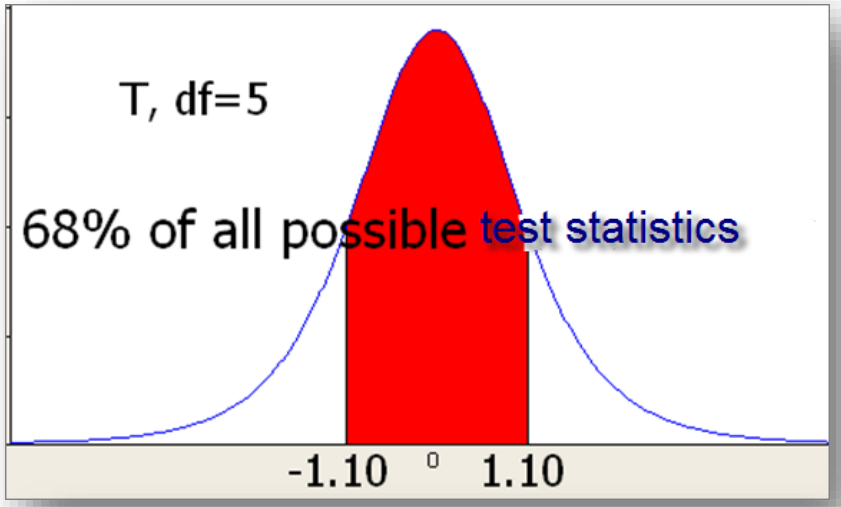
Hypothesis Testing needs a Null

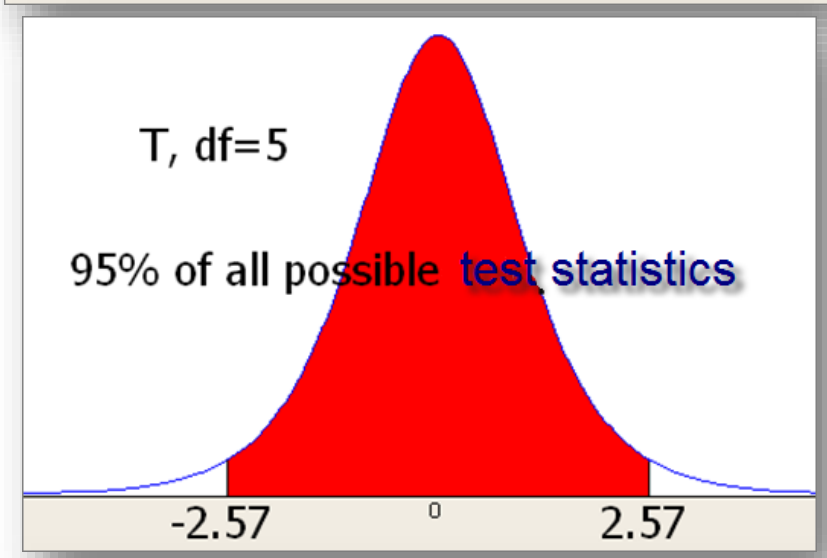
- For hypothesis testing, we follow:
  - Model
  - Compare
  - Ask
- Knowing how sample means behave, we can use this to define a *Null Model*

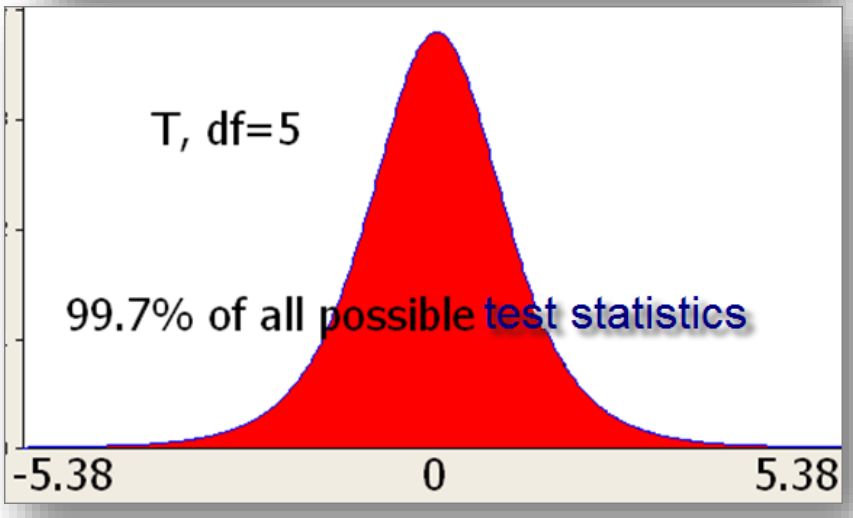


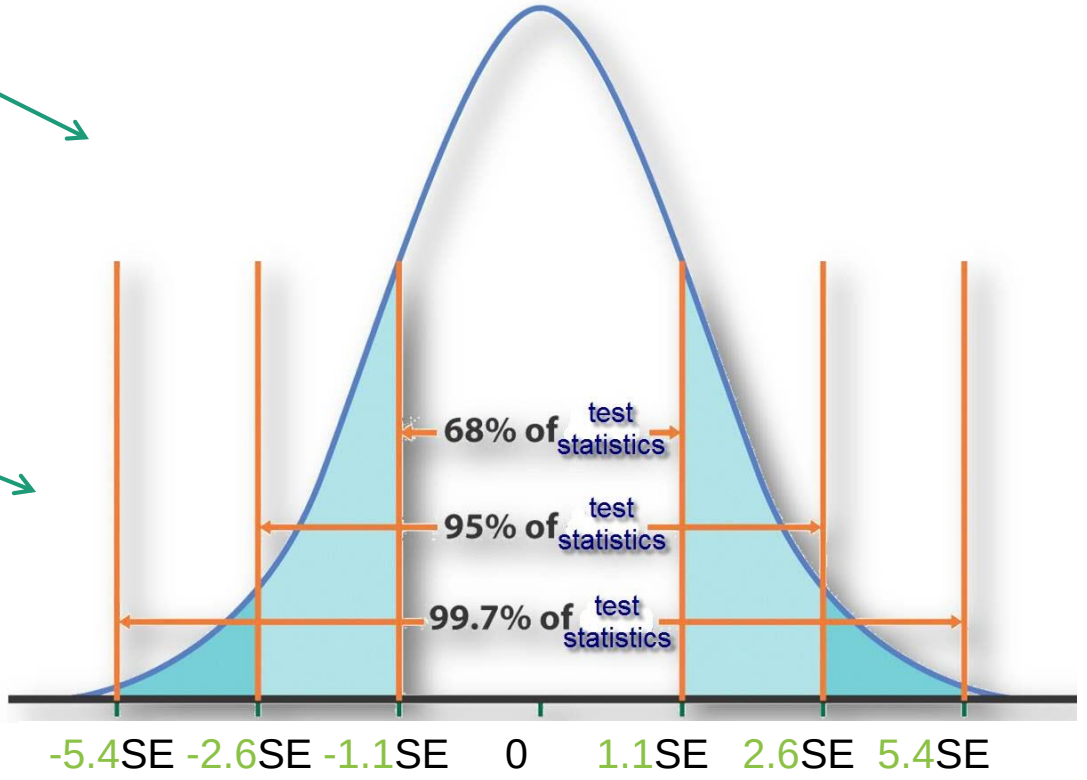
A Two-Sample Example

Small Sample Size



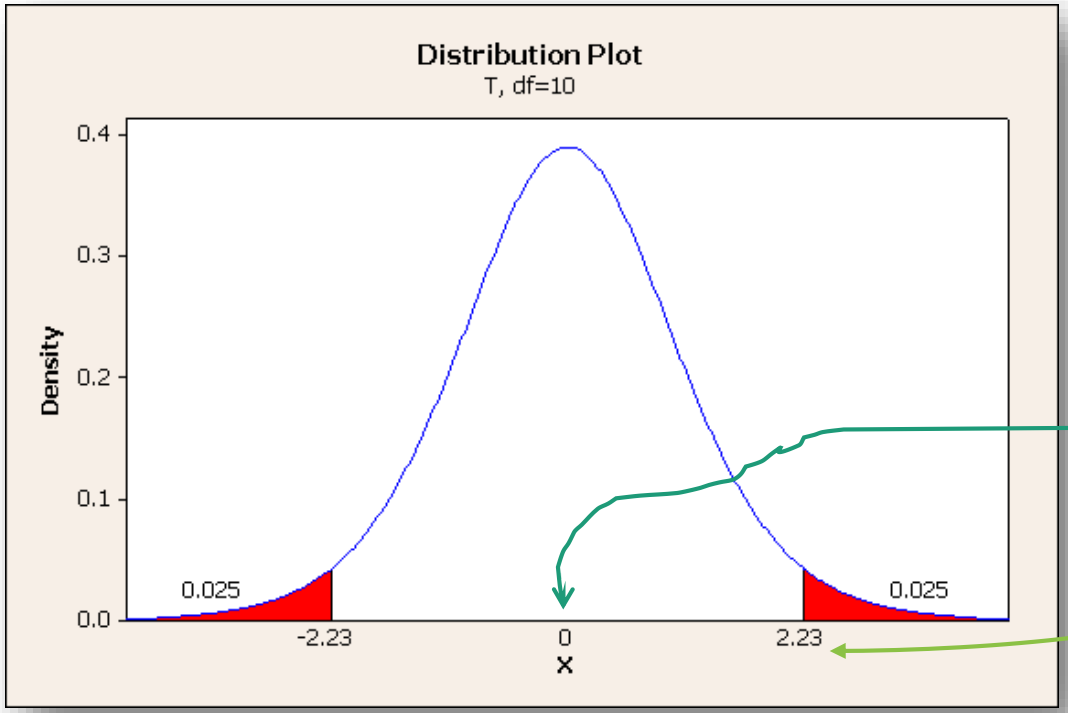






The two-sample t-statistic

$$t_{stat} = \frac{(\text{sample mean\_1} - \text{sample mean\_2})}{\text{standard error}} = \frac{(\bar{x}_2 - \bar{x}_1)}{SE}$$



Allows us to work with null model centered on zero

Allows us to see how many standard errors our difference is from the null

The p-value

- Once we calculate our t-stat from our data, a p-value is also generated that, *in a number*, tells us whether our data was likely or unlikely to be found, **IF** the null is true.
- The p-value is called a conditional probability.
- *On the condition that the null is true*, it's the probability of getting data as different from the null mean (or more different) as we did.

- Small p-values are good evidence against the null
- Large p-values are poor evidence

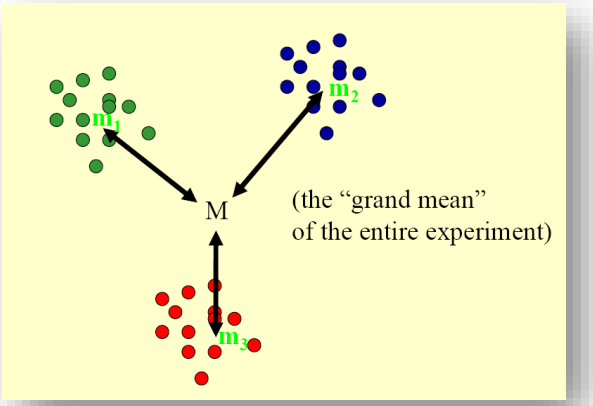
Variance -- *the square of standard deviation* -- has this general form:

$$s^2 = \frac{\sum_{1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\text{Sum of Squares}}{\text{Degrees of Freedom}} = \frac{\text{SS}}{\text{df}} = \text{MS}$$

- Variance is also called a *Mean Square* and abbreviated as MS

One-Way ANOVA

partitions the sources of variability



Total Sum of Squares

SS_{Total}

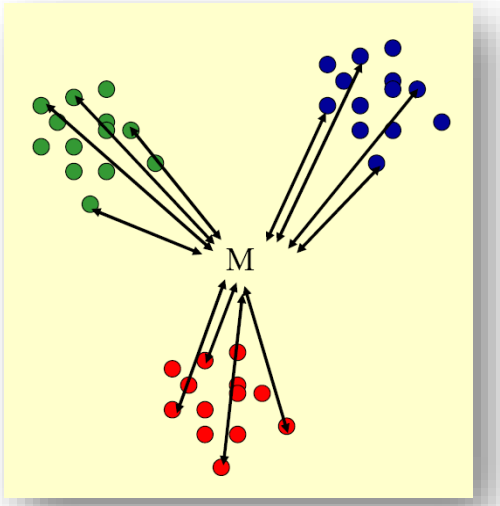
Between (Factor) Sum of Squares

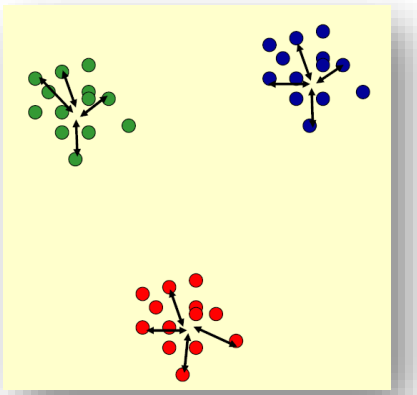
SS_{Factor}

Within (Error) Sum of Squares

SS_{Error}

Find F_{stat}





The natural *F* statistic

- The natural statistic that comes out of separating out these variance is the F-statistic

$$F = \frac{\text{variance}_{\text{between}}}{\text{variance}_{\text{within}}} = \frac{\text{variance}_{\text{treat}}}{\text{variance}_{\text{error}}} = \frac{MS_{treatment}}{MS_{error}} = \frac{signal}{noise}$$

- You can see that as this number gets larger than 1, we can start to detect differences between treatment groups over the noise

ANOVA Summary Table

Source ↓	df	Sums of squares, SS	Mean square, MS (aka variance)	*F*-ratio
Treatment (aka *Between*)				
Error (aka *Within*)				
Total				

EXAMPLE *for media formulation study*

The basic principles of experimental design (Fisher, 1930)

- **Factorial principle**
  - Treatments are generated by combining the levels of factors
- **Randomization**
  - The assignment of treatments to the experimental material, the order in which the runs are to be performed and other aspects of experiments are randomly determined
- **Replication**
  - An independent repeat of each factor combination (experiment)
  - Estimation of experimental error
- **Blocking**
  - Used to reduce the variability induced by nuisance factors

Example: Varieties of Wheat

- One of the earliest published example of a complete, randomized block design was from Sir Ronald Fisher's 1935 book, *The Design of Experiments*
- Goal: compare five varieties of wheat for highest yield
- Design:
  - Treatment: variety of wheat
  - Response: yield in bushels per acre
  - Use blocks

Nuisances

- A nuisance is any possible source of variability other than the conditions you want to compare
  - Anything other than the effects of interest (i.e. signal) that might affect the response
- For example, known differences in the terrain (soil, light, water) will be a nuisance to the design and our ability to “see” a difference

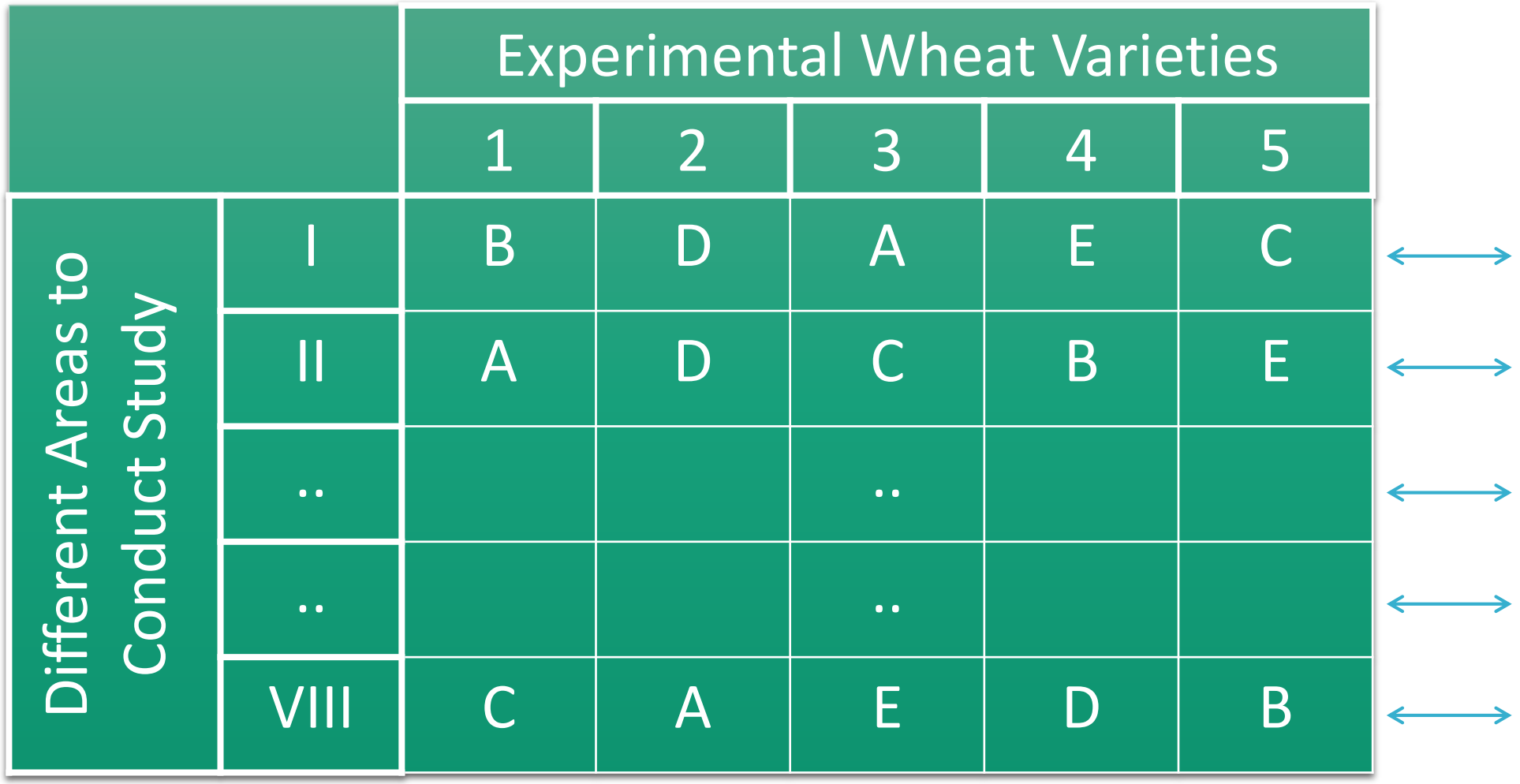
Nuisances

- Randomizing turns a nuisance influence into chance error
  - Random assignment turns possible bias into chance error (e.g. this gets added to our MS_{error} term)
- Blocking turns nuisance influence into a *factor* of the design
  - Sort your material (i.e. experimental units) into subgroups where within each *the nuisance influence is similar* then run a bunch of mini-completely randomized experiments in parallel, one for each group

Wheat example: nuisances

- Weather – some growing seasons better than others
- Land – variation in soil
- Fisher had 8 areas of land to work with
  - Knowing that each piece of land was different – he wanted to **block the influence *between* different areas**
  - He subdivided each area into 5 plots, one for each variety
  - Each area was it's own mini-CR experiment

Fisher's Design



Large variation in nuisance variable(s) (vertically)

Little variation in nuisance variable(s) (horizontally)

ANOVA with Blocks

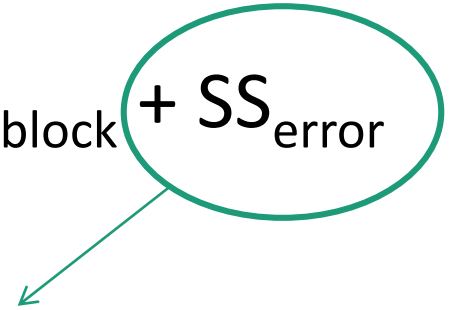
- We take advantage of

$$\text{Total SS} = SS_{\text{treatment}} + SS_{\text{error}}$$

The ability to attribute variability to different sources

- To now become

$$\text{Total SS} = SS_{\text{treatment}} + SS_{\text{block}} + SS_{\text{error}}$$



This is in the denominator of our test statistic; if we can make this smaller with blocks = better design

Source	df	Sums of squares	Mean square	*F*-ratio
Treatment				
Error				
Total				

Our original ANOVA gets a new row added to the table

Source	df	Sums of squares	Mean square	*F*-ratio
Treatment				
Blocks				
Error				
Total				

EXAMPLE *for media formulation study*

Handling influential variables in an experiment

- If you can (and want to), **fix** an influential variable
  - e.g., use only one media formulation, cell strain, process condition
  - Downside?

- If you don't/can't fix an influential variable, **block** its effect
  - e.g., block *the influence of the variable*
  - Downside?

- If you can neither fix nor block a variable, **randomize it**
  - e.g. randomize to deal with unknown factors

  >> "Block what you can, randomize the rest"

ANOVA and Linear Regression

- Simple linear regression is a one-way ANOVA
  - $y = mx + b$
  - x is the single factor (with some number of levels) describing the response, y
- Multiple linear regression includes more than one factor
  - $y = m_1x_1 + m_2x_2 + \ldots + b$
  - Each x is a factor (with some number of levels) describing the response, y

- *Different sides of the same coin…*

ANOVA and the regression

- r^2 is one of the more abstract concepts in regression
- This value comes from an ANOVA analysis
  - $SS_{Total} = SS_{Regression} + SS_{Error}$

$$r^2 = \frac{SSR}{SST} = \frac{\text{sum of } (y_{\text{Predicted}} - \bar{y})^2}{\text{sum of } (y_{\text{Observed}} - \bar{y})^2}$$

Introduction to Design of Experiments

Section Three

Definition of DoE

Statistical design of experiments:

- The process of planning the experiment so that appropriate data that can be analyzed by statistical methods will be collected resulting in valid objective conclusions. [*D. C. Montgomery*]

- DoE is a structured, organized method for determining the relationships among factors affecting a process and its output. [*ICH Q8*]

Strategy of experimentation: OFAT vs. DOE

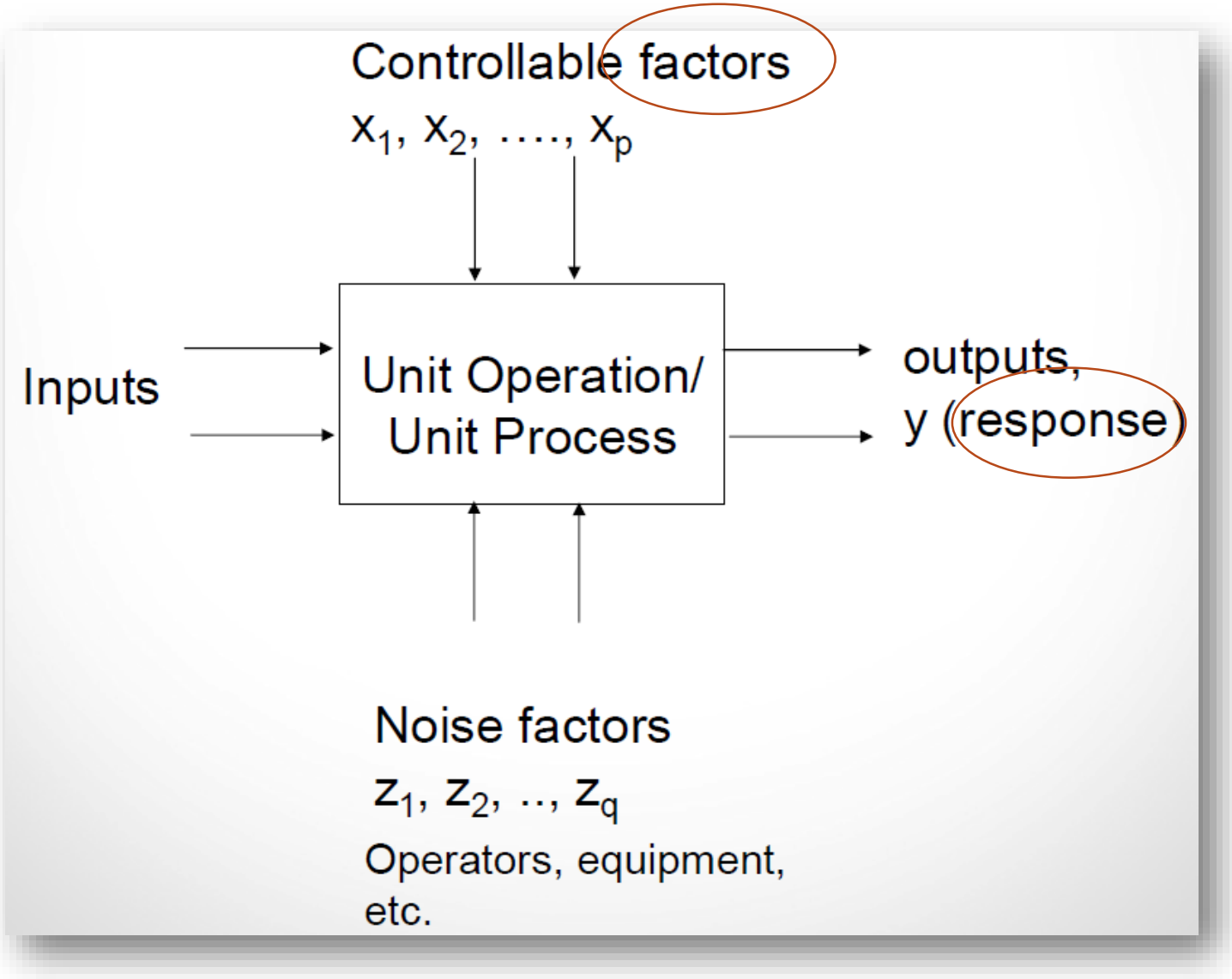
Traditional approach to experimentation

- Study one variable (factor) at a time (OFAT) holding all other variables constant;
- Simple process, but doesn't account for interactions;
- It is inefficient.

Factorial design or statistically designed experiments

- Study multiple factors changing at once;
- Accounts for interactions between variables;
- Maximize information with minimum runs.

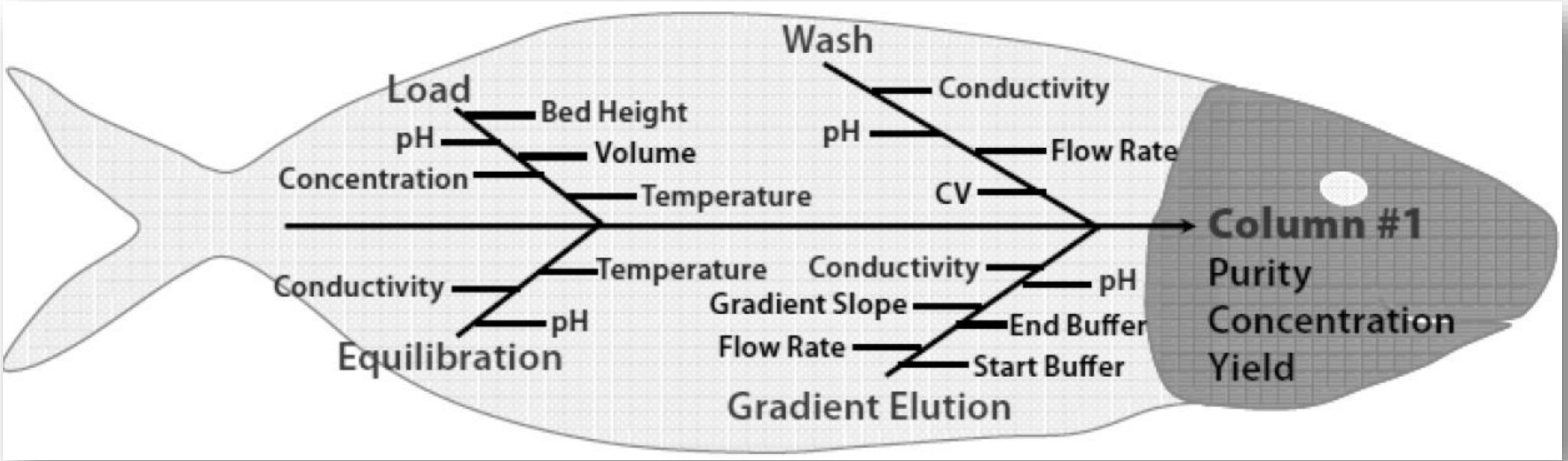
Typical unit operation or process



Examples of *factors* and *responses* in cell culture

- Controllable factors, x_i
  - Temperature
  - pH
  - Agitation rate
  - Dissolved oxygen
  - Medium components
  - Feed type and rate
- Responses, y_i
  - Product concentration
  - Cell viability
  - Product characteristics (glycosylation, ..)

Factors and responses for column chromatography



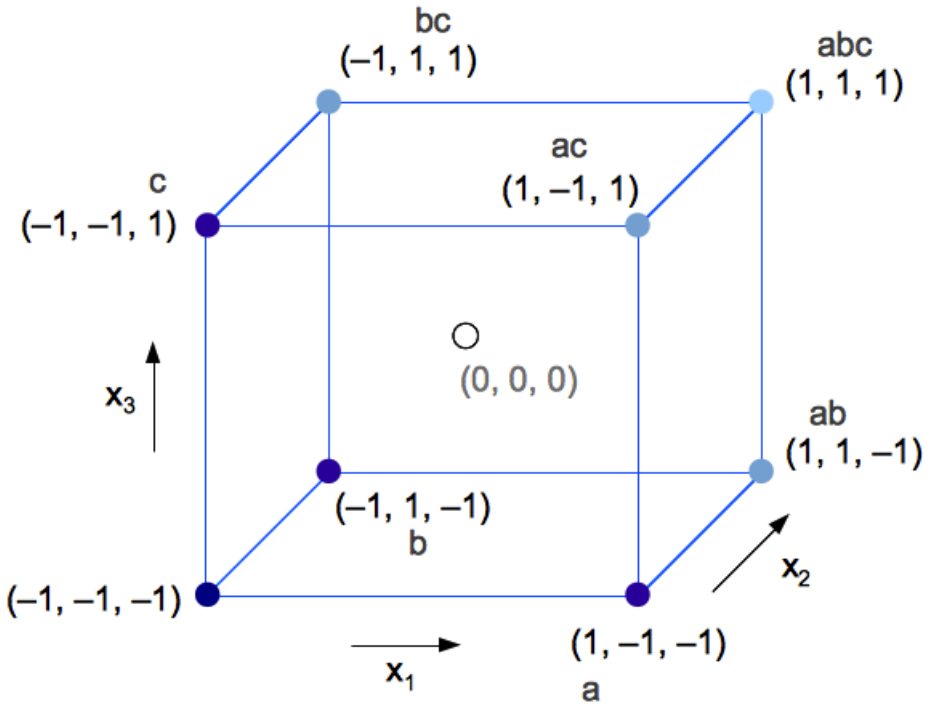
Phases of a DoE process: planning, conducting and analyzing an experiment

1. Statement of problem
2. Choice of factors, levels, and ranges
3. Selection of the response variable(s)
4. Choice of design
5. Conducting the experiment
6. Statistical analysis
7. Drawing conclusions, recommendations

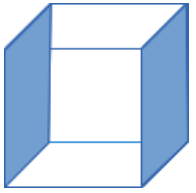
DoE helps only with points 4 and 6!

The most common 2^k full factorial design

The classic 2^3 full factorial (2-level 3 factors) design graphically:



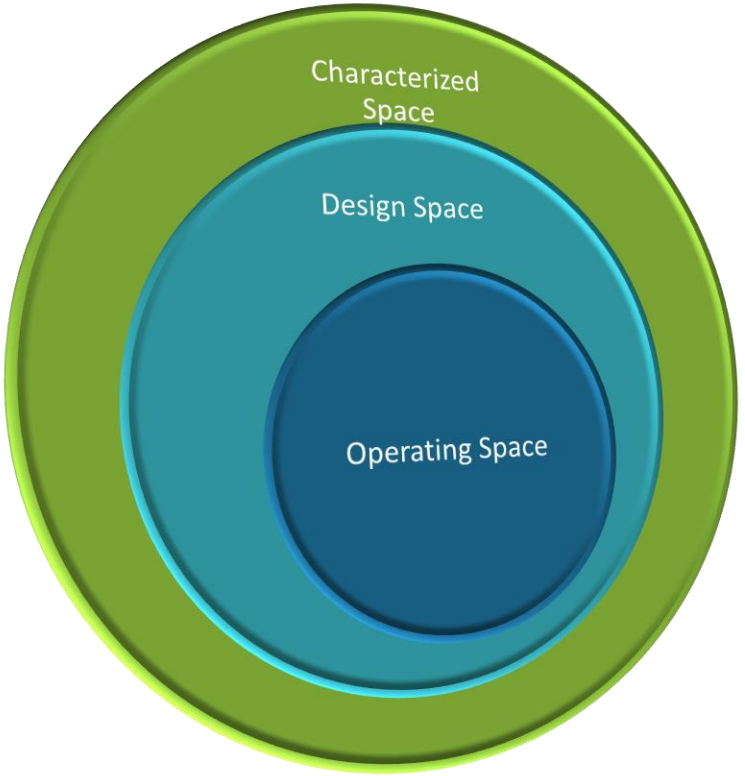
The points involved in the sample calculations of the main effects of *A* (X_1):



and the interaction of *A* & *C* (X_1X_3):

DoE objectives and process spaces

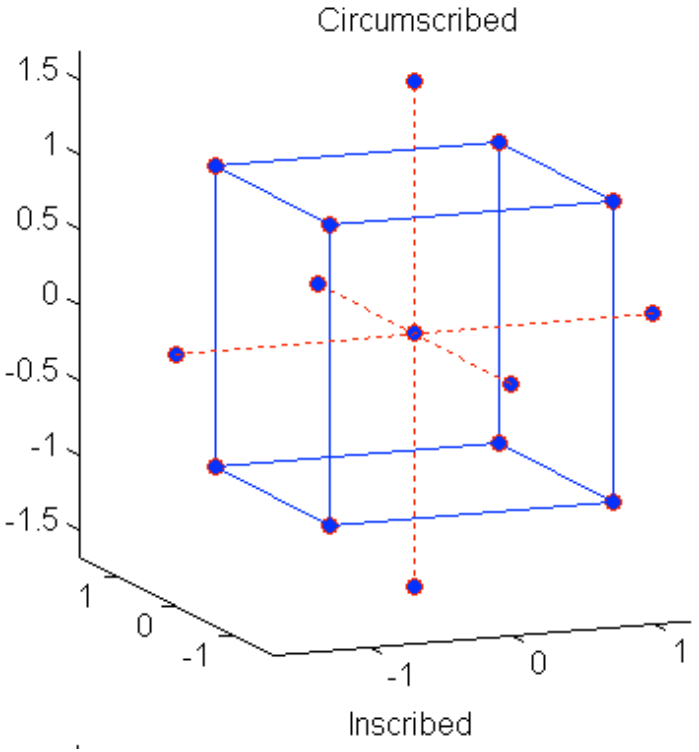
- Screening/Characterization
  - Which factors are important?
  - What are the appropriate ranges for these vital factors?
- Optimization
  - Detailed quantification of the effect of the vital factors
  - What are the optimal ranges for these factors?
- Robustness testing
  - Verify that process is robust to small variations in the input parameters



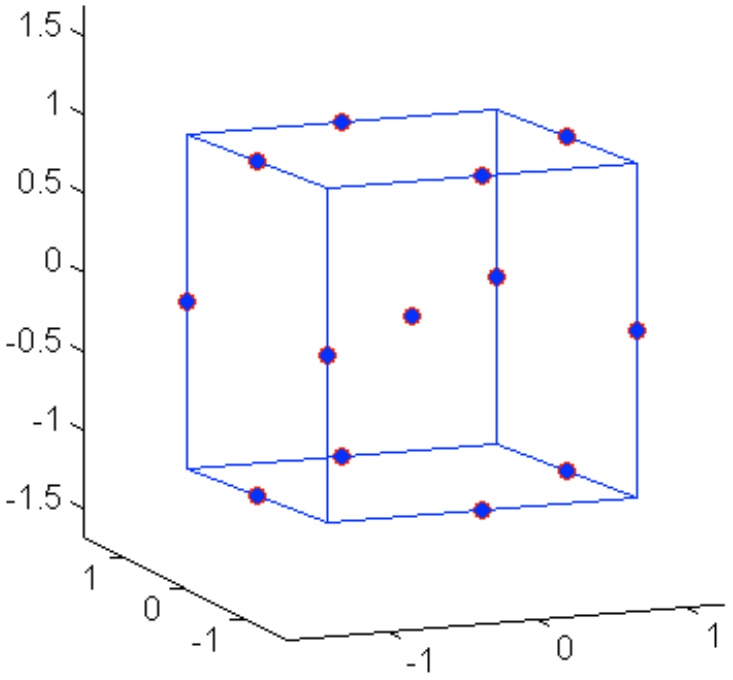
There are numerous other designs

Can find them (and their purpose) in texts and generate them using statistics packages.

Two images from Matlab:



A circumscribed form of a central **composite** design (CCDs), a.k.a. Box-Wilson designs, with center and star points.



A Box-Behnken design. Note that it avoids the corners of the design space—maybe a good thing if they are extreme conditions.

A catalogue of designs

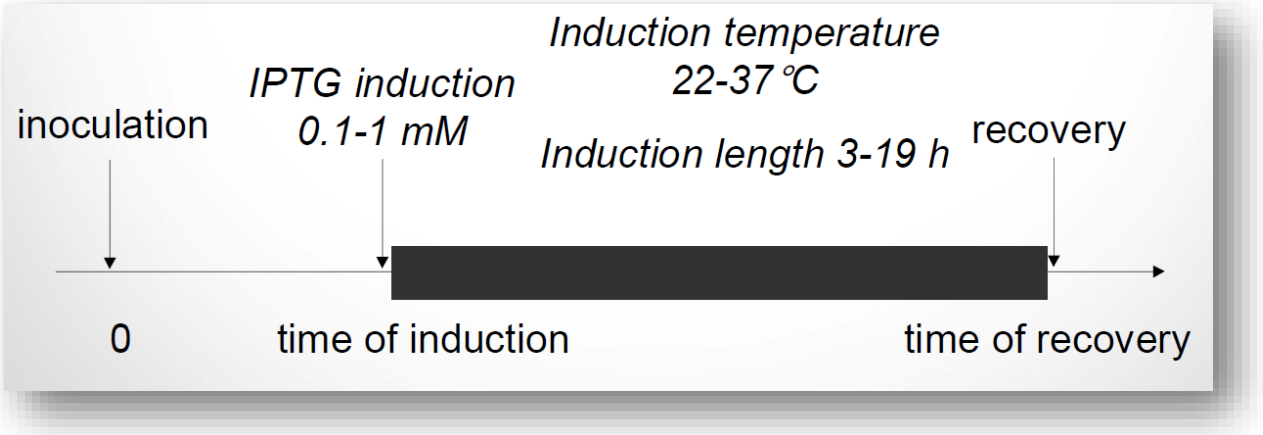
Design	Use
Full Factorial	Characterization
Fractional Factorial	Screening
Plackett-Burman	Screening
Central Composite	Optimization
Box-Behnken	Optimization
Mixture	For mixtures (factors are compositions: ex, $x_1+x_2+x_3=1$)

Design Selection Guideline

Number of Factors	Comparative Objective	Screening Objective	Response Surface Objective
1	1-factor completely randomized design	–	–
2 - 4	Randomized block design	Full or fractional factorial	Central composite or Box-Behnken
5 or more	Randomized block design	Fractional factorial or Plackett-Burman	Screen first to reduce number of factors

A 2^3 replicated factorial design: GFP expression by *E. coli* in baffled shake flasks

- *Medium:*
  - Bacto Yeast Extract - 25 g/L; Tryptic Soy Broth - 15 g/L; NH4Cl - 1 g/L; Na2HPO4 - 6 g/L; KH2PO4 - 3 g/L; Glucose - 10 g/L.
- *Culture conditions:*
  - 250-mL baffled shake flasks, 25-mL culture volume, agitation speed 400 rpm, growth temperature 37°C.



Defining the factors and their levels

Several factors affect GFP expression:

- Induction temperature
  - generally 37°C or lower. During induction the temperature can be decreased with respect to the growth phase;
- Induction length
  - three hours allows to recover the cells the same day of inoculation; 19 h corresponds to an overnight;
- Inducer concentration
  - generally the range 0.1-1 mM is used. Using a small quantity of inducer saves money.

Factor Levels	Low (-1)	High (+1)
Induction temperature (A)	23 ℃	37 ℃
Induction length (B)	3 h	19 h
Inducer concentration (C)	0.1 mM	1 mM

Choosing the design: a 2^3 full factorial design

St. ord.	Run	Coded Factors			Factor Levels		
		A	B	C		Low (-1)	High (+1)
1	(1)	-1	-1	-1	A (℃)	22	37
2	a	+1	-1	-1	B (h)	3	18
3	b	-1	+1	-1	C (mM)	0.1	1
4	ab	+1	+1	-1			
5	c	-1	-1	+1			
6	ac	+1	-1	+1			
7	bc	-1	+1	+1			
8	abc	+1	+1	+1			

**Replicated twice*

Run order is the randomized standard order

Running the experiment

RunOrder	StdOrder	Temperature ℃	Time h	Inducer mM	GFP conc. mg/mL
1	5	22	3	1	0.25
2	16	37	18	1	0.74
3	10	37	3	0.1	0.67
4	12	37	18	0.1	0.88
5	14	37	3	1	0.60
6	4	37	18	0.1	0.87
7	9	22	3	0.1	0.21
8	15	22	18	1	1.82
9	2	37	3	0.1	0.45
10	3	22	18	0.1	1.95
11	7	22	18	1	1.80
12	11	22	18	0.1	1.71
13	6	37	3	1	0.43
14	13	22	3	1	0.41
15	1	22	3	0.1	0.36
16	8	37	18	1	0.92

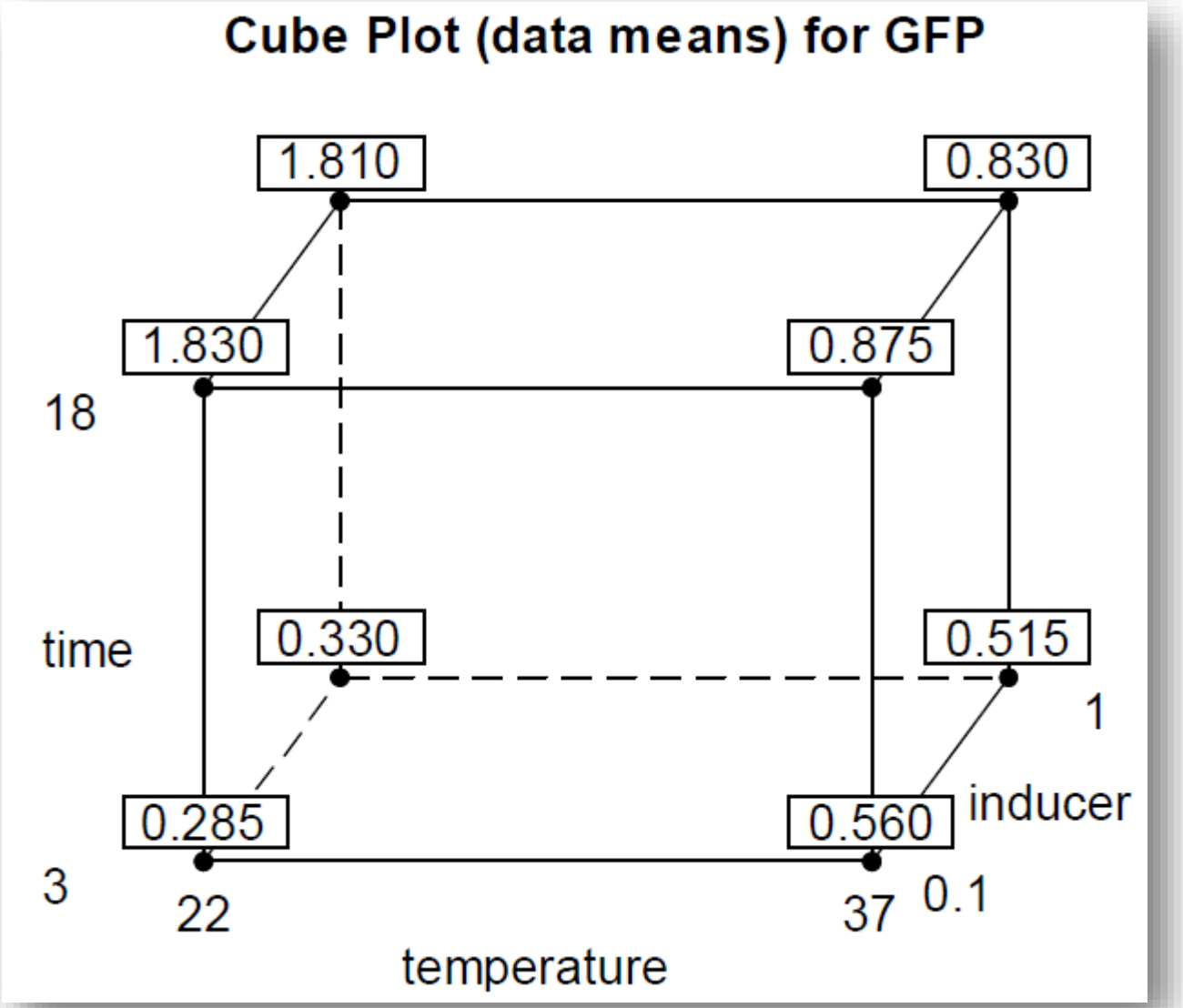
Also OD_{600} data are available

The experiment is replicated once (n=2).

Sometimes we say it is replicated two times to mean the same.

Experiments are carried out according to the run order. Several aspects of the experiment are randomized (inoculation, induction, position in the shaker, etc.)

Cube Plot


Cube Plot (data means) for GFP
1.810
0.830
1.830
0.875
18
time
0.330
0.515
1
0.285
0.560
inducer
3
22
37
0.1
temperature

ANOVA – Minitab Output 1

Analysis of Variance for GFP, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
(A) temperature	1	0.54391	0.54391	0.54391	40.31	0.000
(B) time	1	3.33976	3.33976	3.33976	247.50	0.000
(C) inducer	1	0.00106	0.00106	0.00106	0.08	0.787
(AB) temperature*time	1	1.43401	1.43401	1.43401	106.27	0.000
(AC) temperature*inducer	1	0.00331	0.00331	0.00331	0.25	0.634
(BC) time*inducer	1	0.00106	0.00106	0.00106	0.08	0.787
(ABC) temp.*time*inducer	1	0.00106	0.00106	0.00106	0.08	0.787
Error	8	0.10795	0.10795	0.01349		
Total	15	5.43209				

S = 0.116163 R-Sq = 98.01% R-Sq(adj) = 96.27%

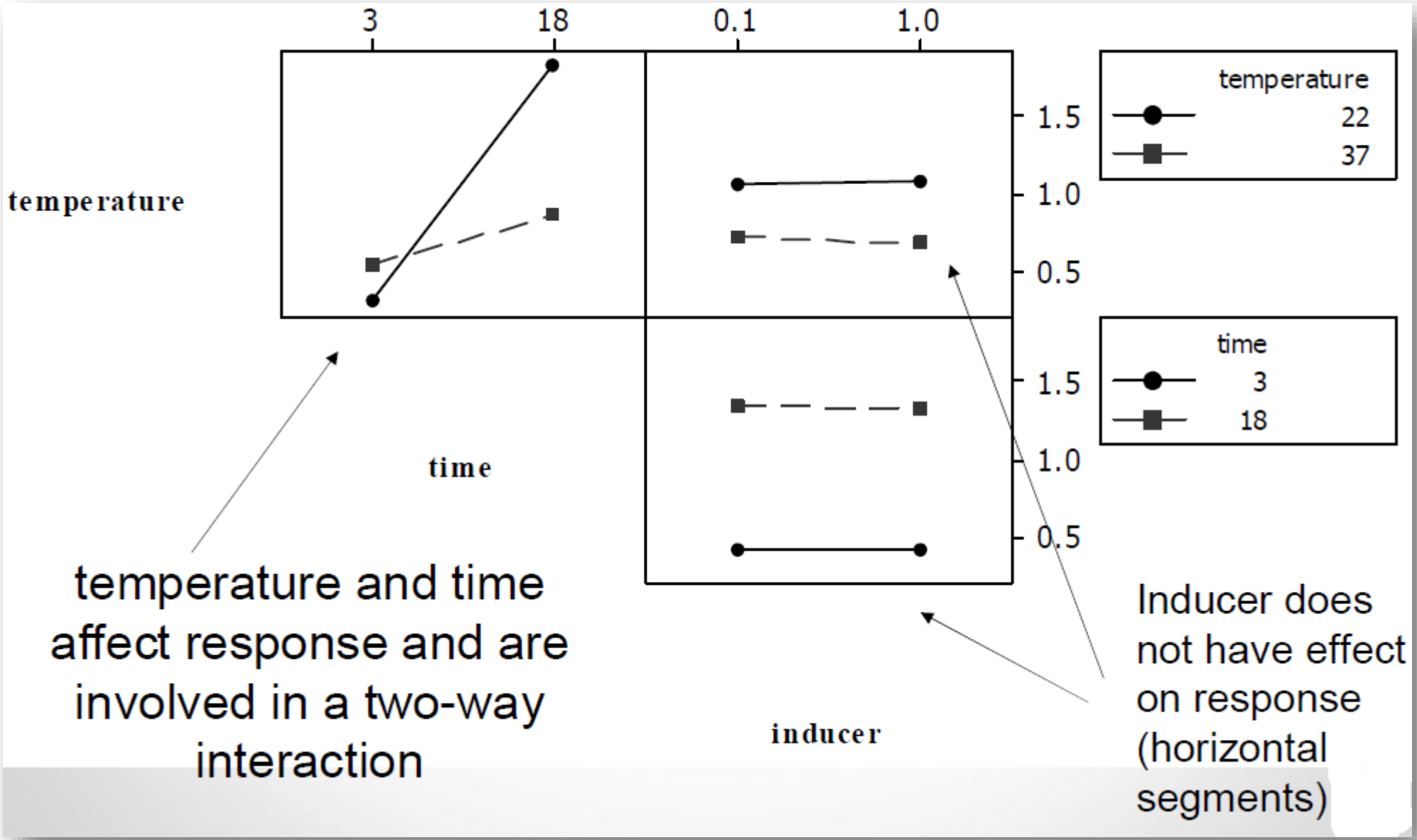
Effects, regression coefficients – Minitab Output 2

Estimated Effects and Coefficients for GFP (coded units)

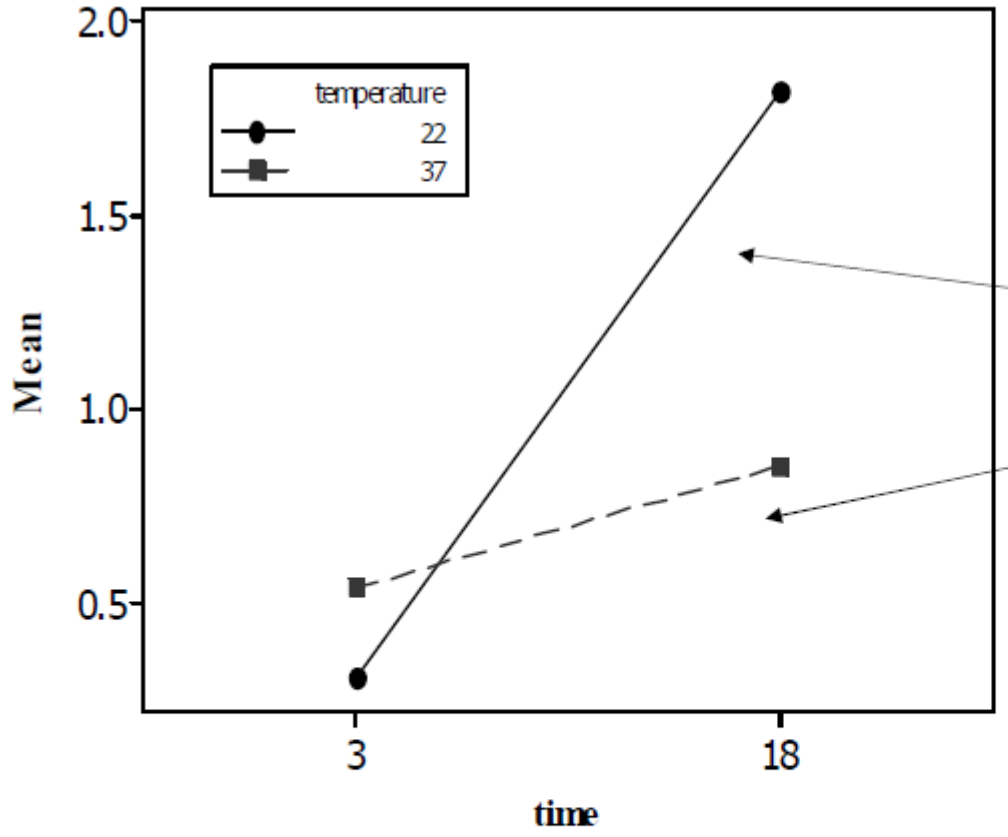
Term	Effect	Coef	SE Coef	T	P
Constant		0.8794	0.02904	30.28	0.000
temperature	-0.3687	-0.1844	0.02904	-6.35	0.000
time	0.9138	0.4569	0.02904	15.73	0.000
inducer	-0.0162	-0.0081	0.02904	-0.28	0.787
temperature*time	-0.5988	-0.2994	0.02904	-10.31	0.000
temperature*inducer	-0.0288	-0.0144	0.02904	-0.49	0.634
time*inducer	-0.0163	-0.0081	0.02904	-0.28	0.787
temperature*time*inducer	0.0162	0.0081	0.02904	0.28	0.787

S = 0.116163 PRESS = 0.4318
R-Sq = 98.01% R-Sq(pred) = 92.05% R-Sq(adj) = 96.27%

Interpreting results: interaction plot



Temperature x time interaction plot



Time has a
high effect at 22°C
and
a small effect at 37°C

At 3 h of induction, temperature has a positive effect while at 18 h of induction temperature has a negative effect.

The best condition is 18 h of induction at 22°C at whatever level of inducer (low or high).

Interpreting results

- The main effect of the inducer concentration (factor C) and all its interactions (AC, BC, ABC) are not significant.
- When we changed the level of C in the experiment it was like if we were replicating a treatment (for example, treatment abc and treatment ab are considered replicates).
- We would therefore work with a reduced model that explains GFP titer…

Effects, regression coefficients – Reduced model

Estimated Effects and Coefficients for GFP (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		0.8794	0.02441	36.02	0.000
temperature	-0.3687	-0.1844	0.02441	-7.55	0.000
time	0.9138	0.4569	0.02441	18.71	0.000
temperature*time	-0.5988	-0.2994	0.02441	-12.26	0.000

S = 0.0976495 PRESS = 0.203422
R-Sq = 97.89% R-Sq(pred) = 96.26% R-Sq(adj) = 97.37%

- Regression equation in coded units:

$$\hat{y} = 0.879 - 0.184x_1 + 0.457x_2 - 0.299x_1x_2$$