AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

A Primer on Scientific Test & Analysis Techniques (STAT)

A Tutorial at
ITEA's TIW and DETE Workshop
May 12, 2015
Las Vegas, NV

15-STATPRIM-5A

Mark J. Kiemele, Ph.D.
President and Co-Founder
Air Academy Associates

Office: 719-531-0777
Cell: 719-337-0357
mkiemele@airacad.com
www.airacad.com

Introductions

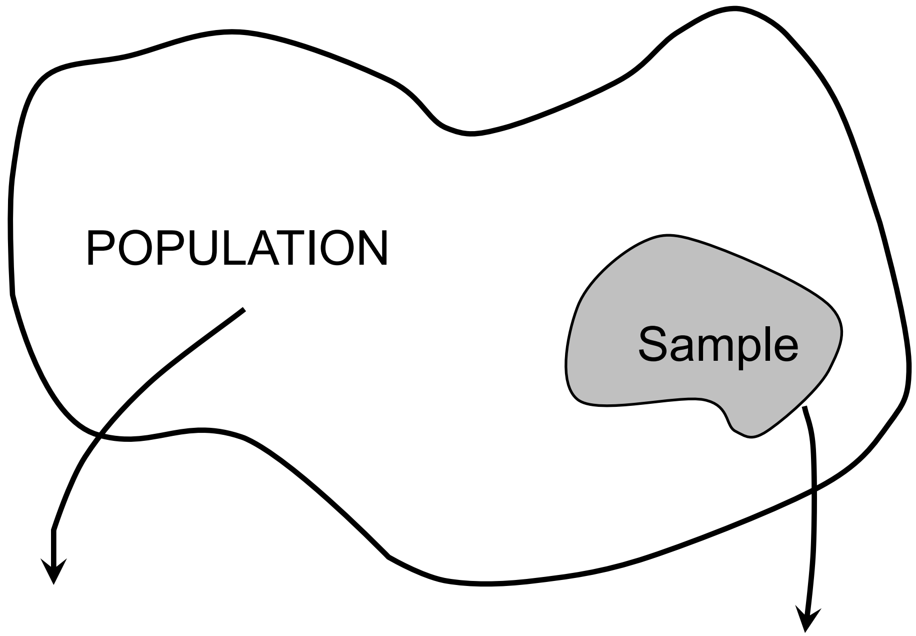
- **Name**
- **Organization**
- **Job Title/Responsibilities**
- **Experience in test and analysis techniques, such as Measurement System Analysis, Design of Experiments (DOE), etc.**

Agenda

- Some Basic Definitions and Terms
- Hypothesis Testing (Confidence and Power)
- Measurement System Analysis (MSA)
- Demonstration of a MSA
- Design of Experiments (DOE): a Modern Approach to Testing
- Examples and Demonstration of a Modeling DOE
- Best Practices for “operationalizing” these techniques

Air Academy Associates
Simplify, Perfect, Innovate

Basic Statistics Review



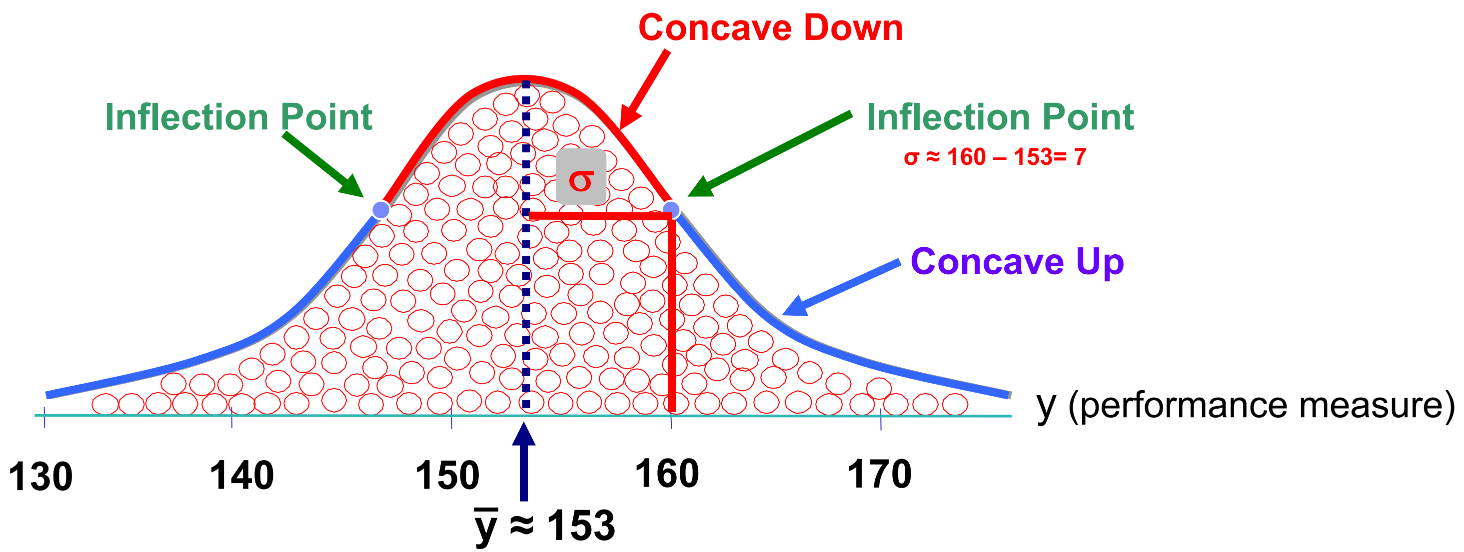
Population Parameters		Sample Statistics
μ	**Mean**	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
σ^2	**Variance**	$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$
σ	**Standard Deviation**	$s = \sqrt{s^2}$

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Graphical Meaning of μ ($\bar{y}$) and σ (s)

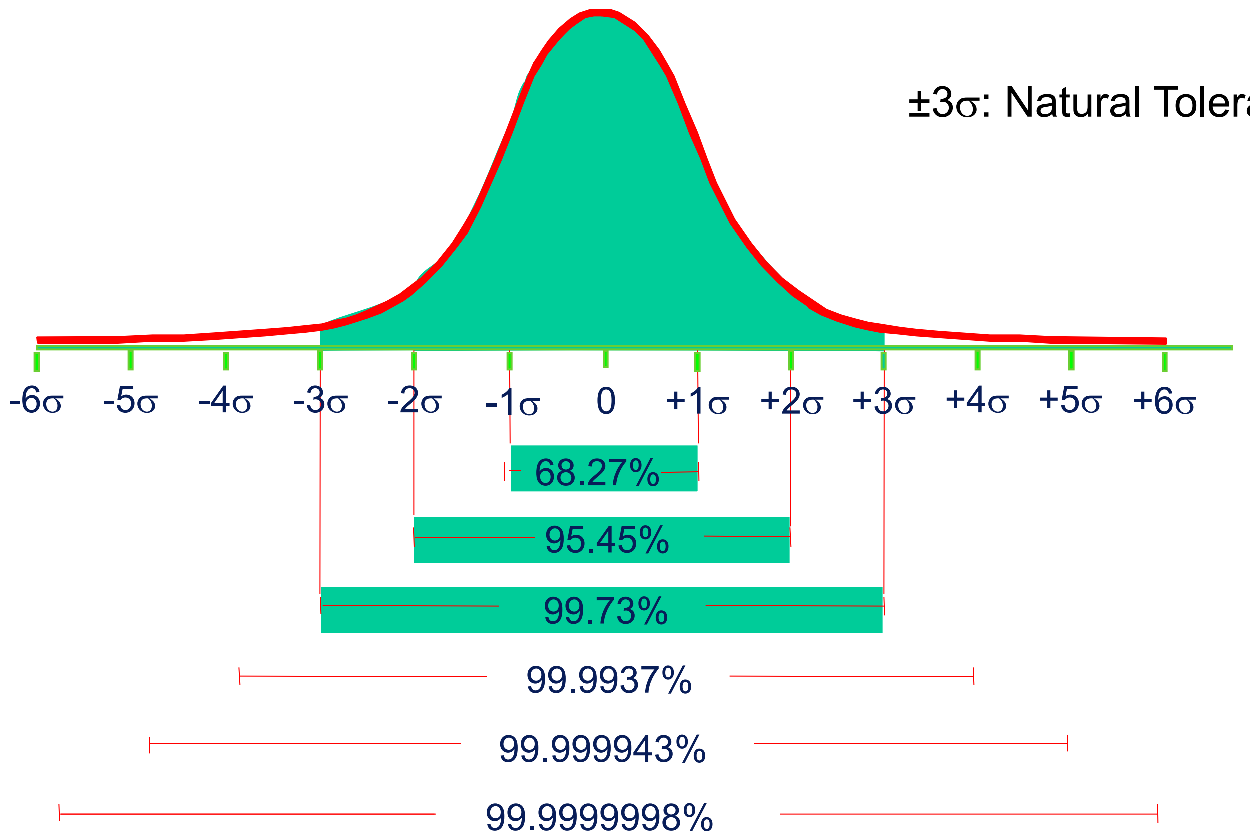
$\bar{y}$ = Average = Mean = Balance Point **(location)**

σ = Standard Deviation **(shape)**



σ ≈ average distance of points from the centerline

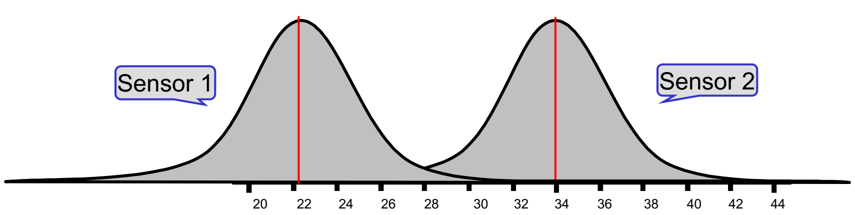
Graphical View of Variation



Typical Areas under the Normal Curve

Hypothesis Testing

- A method for looking at data and comparing results
  - Method A vs. Method B
  - Material 1 vs. Material 2
  - Sensor 1 vs. Sensor 2
- Helps us make good decisions and not get fooled by random variation:
  - "Is a difference we see REAL, or is it just random variation and no real difference exists at all?"
- We set up 2 hypotheses
  - H_0 is called the null hypothesis (no change, no difference)
  - H_1 is called the alternate hypothesis
  - Example: $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$



  - Based on the data we collect, we must decide in favor of either H_0 or H_1. Which does the evidence support?

Nature of Hypothesis Testing

H_0: Defendant is Innocent

H_1: Defendant is Guilty

Since verdicts are arrived at with less than 100% certainty, either conclusion has some probability of error. Consider the following table.

		True State of Nature	
		H_0	H_1
Conclusion Drawn	H_0	Conclusion is Correct	Conclusion results in a Type II error
	H_1	Conclusion results in a Type I error	Conclusion is Correct

Type I or Type II Error Occurs if Conclusion Not Correct

The probability of committing a Type I error is defined as α ($0 \le \alpha \le 1$) and the probability of committing a Type II error is β ($0 \le \beta \le 1$). The most critical decision error is usually a Type I error.

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Example: 2-Sample Hypothesis Test
(1 Factor at 2 levels)

Target Acquisition Time (in seconds)

Sensor	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	$\bar{y}$	s
1	2.8	3.6	6.1	4.2	5.2	4.0	6.3	5.5	4.5	4.6889	1.17
2	7.0	4.1	5.7	6.4	7.3	4.7	6.6	5.9	5.1	5.8667	1.08

- In a target kill chain process, a team suspects that one of the important contributors to destroying the target is the amount of time it takes a sensor to acquire the target. The team further suspects that there might be a significant difference in the average amount of time it takes two different sensors to acquire the target.

- The team decides to conduct a test, to see if "sensor type" really is an important factor. In other words, is there a significant difference in the average target acquisition times between the first and second sensors?

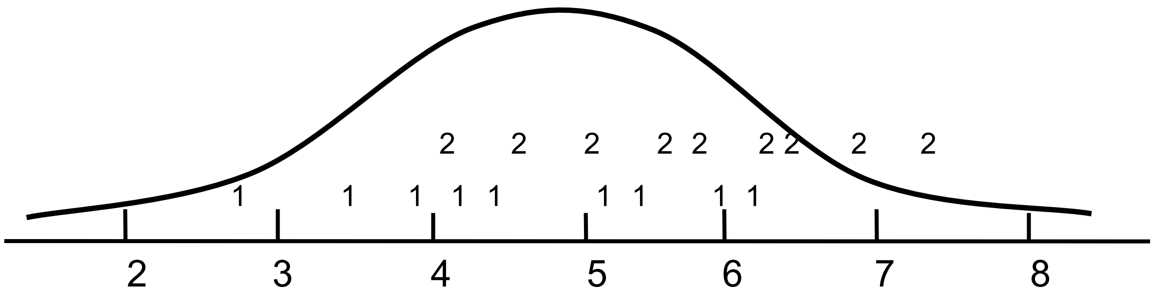
- A random sample of 9 data points for each of the two sensors, shown above, was collected to help answer this question.

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Graphical Interpretation of the Hypothesis Test

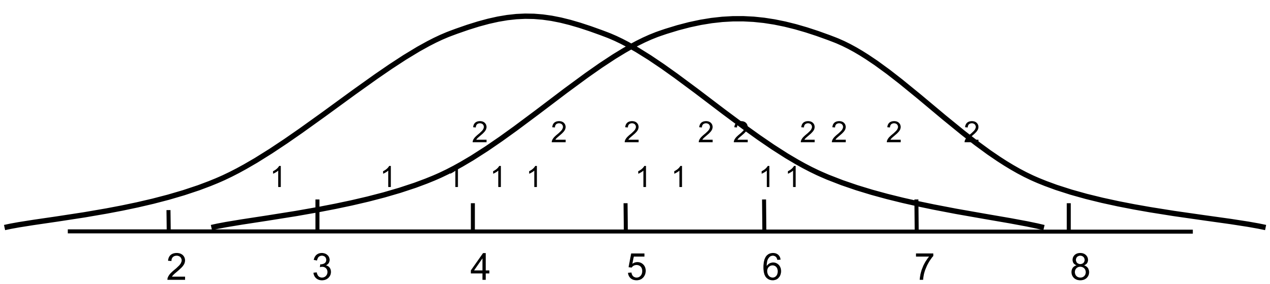
$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$



$H_0: \mu_1 = \mu_2$

versus



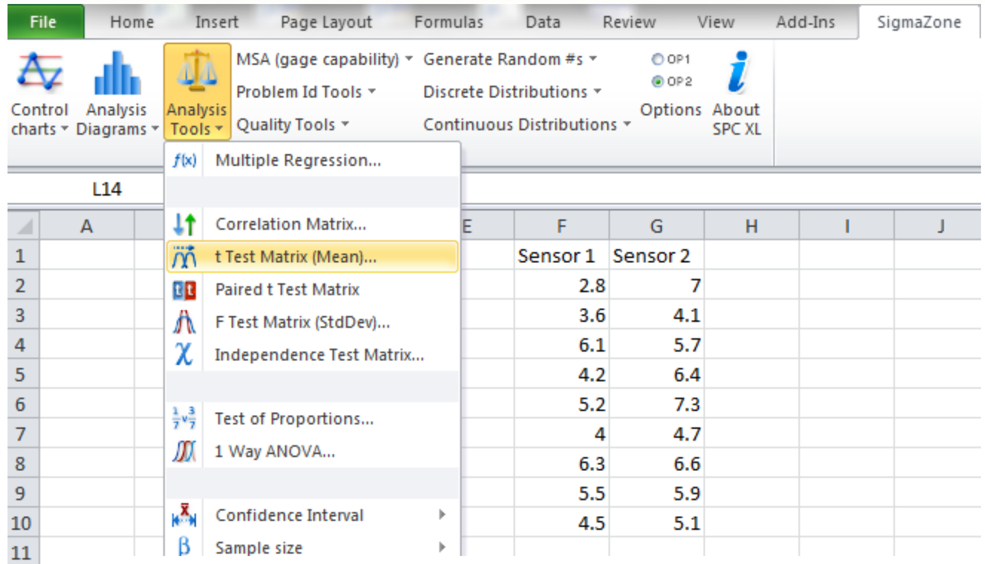
$H_1: \mu_1 \neq \mu_2$

Testing for a Difference in Average

- A formal statistical test for detecting a shift in average is the t-test.
- p-values come from the data and indicate the probability of making a Type I error.
- Rule of Thumb (for significance):
  - If p-value < .05, a highly statistically significant conclusion that H_1 is true
  - If .05 < p-value < .10, a moderately statistically significant conclusion that H_1 is true
  - If p-value > .10, then the result is considered insignificant and we fail to reject the null hypothesis (H_0)
  - (1 – p-value) • 100% is our **percent confidence** that H_1 is true

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

t-test Result on Sensor Example



t-Test Result		
Hypothesis Tested:	H0: Sensor 1 Mean = Sensor 2 Mean H1: Sensor 1 Mean not equal to Sensor 2 Mean	
	p-value (probability of Type I Error)	0.041
	Confidence that Sensor 1 Mean not equal to Sensor 2 Mean	95.9%

Summary Statistics		
	Sensor 1	Sensor 2
Mean	4.6889	5.8667
StDev	1.1731	1.0759
Count	9	9

We are more than 95% confident that H_1: μ1 ≠ μ2 is the correct conclusion. But this says nothing about how different μ1 and μ2 are. **Enter Power!!**

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Determining Sample Size
(for a given α, β, Δ, and σ)

- In the previous example, the team took 9 data points from each sensor, without considering power, and found a statistically significant result. Suppose that, before the test, the team wanted to be able **to detect a change in mean as small as 1 second** between the two sensors' average target acquisition times. What sample size should they have taken?
- Suppose **α = .05**. That is, they want at least 95% confidence.
- Suppose the team wants to have **80% power** in the test. That is, $1 - \beta = .80$, the probability of detecting a change in means between the two sensors as small as $\Delta = 1$ second.
- We still need an **estimate of σ**. Suppose from historical data that the team estimates the standard deviation as 1.1 seconds for both sensors.

Sample Size Calculation for Dual Sided 2 Sample Test (Mean)

User defined parameters	
Estimated Standard Deviation (population 1)	1.1
Estimated Standard Deviation (population 2)	1.1
Size of Difference in Means	1
Desired Power of Test	80.00%
Desired Confidence Level	95.00%
Results	
Sample Size	19

beta risk = 20.00%

alpha risk = 5.00%

Two Types of Risk in Evaluating the Result of a Test

- α Risk = **P(false detection)** means we falsely concluded that a factor is important
  - **P(false detection) = p-value** (software calculates this value from the data)
  - **Confidence = [1 – p-value] x 100%**
  - Rule of Thumb (ROT) for "highly significant" result: Confidence ≥ 95%

- β Risk = **P(missed detection)** means we failed to detect something important
  - **Power = [1 – P(missed detection)] x 100%**
  - Rule of Thumb (ROT) for sufficient power: Power ≥ 75%
  - A Priori (prior to the test) power calculations are good for test planning purposes, and sample size is the way we can control the power of the test.

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Measurement System Analysis Exercise

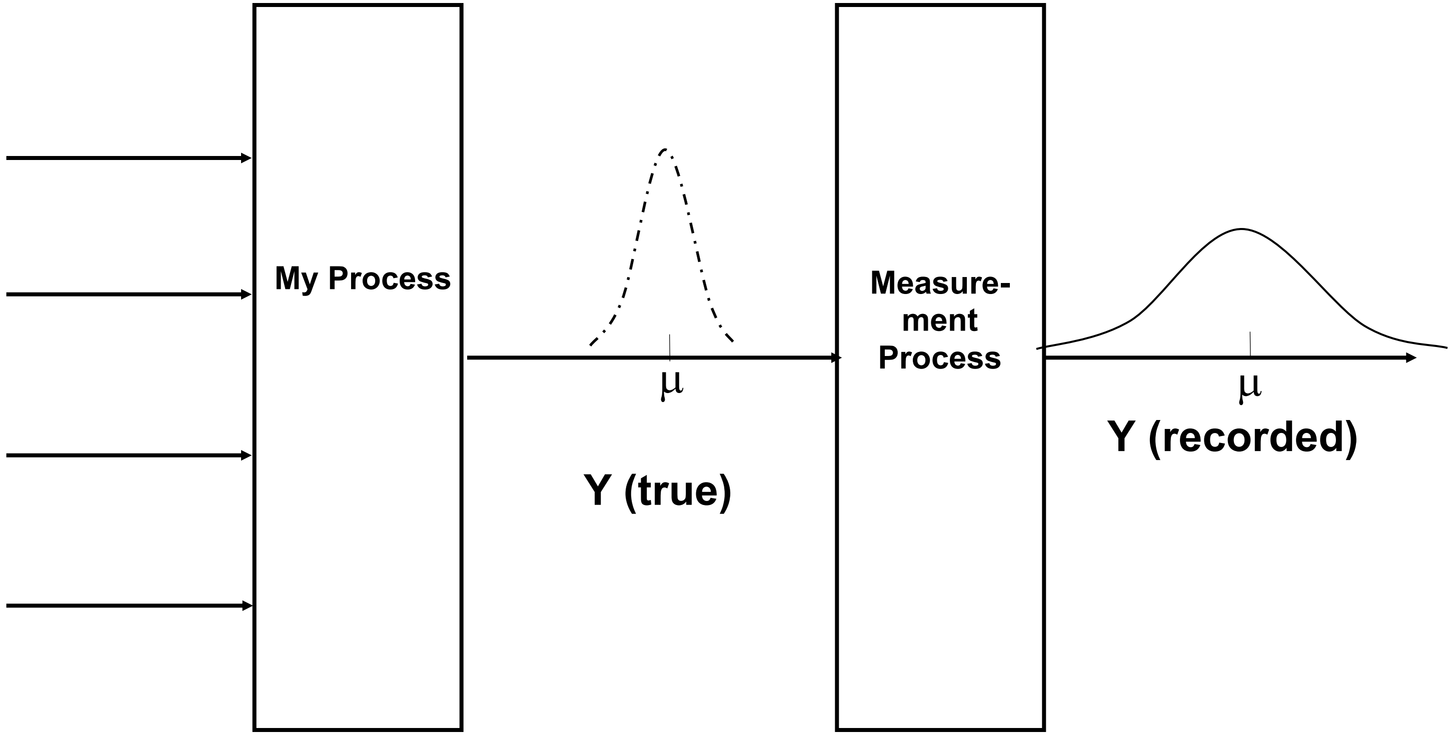
Statapult® Catapult

Data Collection Template

Shot #	Operator 1	Operator 2	Operator 3
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Air Academy Associates
Simplify, Perfect, Innovate

Measurement is a Process



Where **DOES** the variation that we see in Y come from? Is it from the process itself or the measurement system? Which one of those two variations is stronger?

Is the measurement system –

Accurate?

Precise?

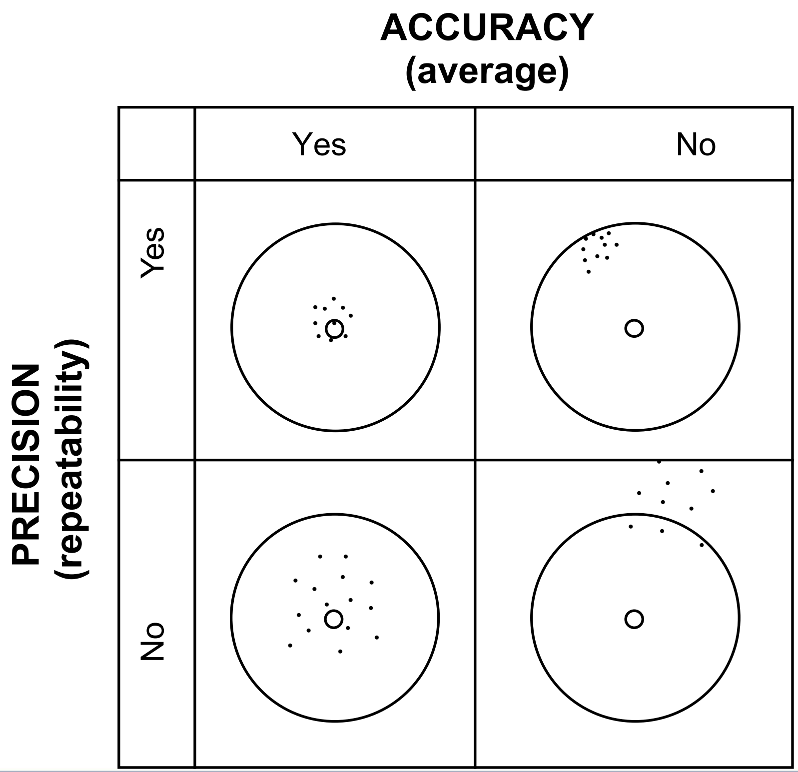
Stable?

Desired Properties of a Measurement System

Accuracy - Ability to produce an average measured value which agrees with the true value or standard

Precision - Ability to repeatedly measure the same product, transaction or service and obtain the same results

Stability - Ability to repeatedly measure the same product, transaction or service over time and obtain the same average measured value

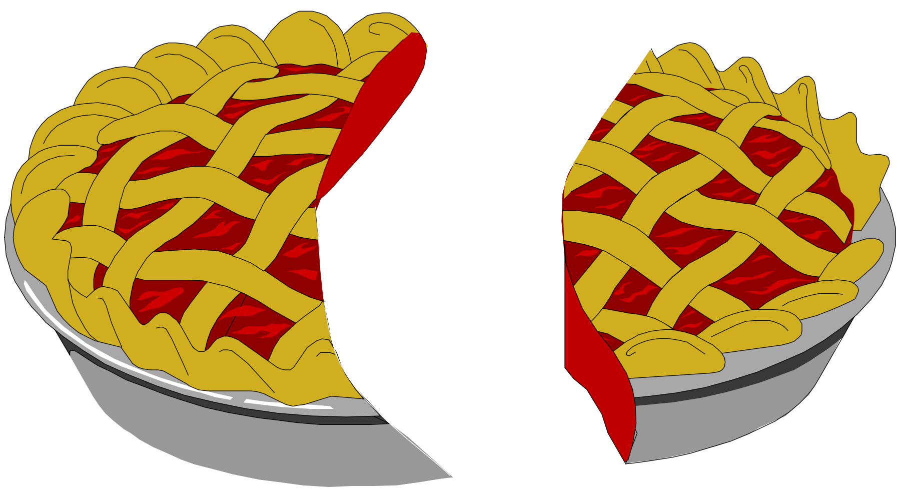


Air Academy Associates
Simplify, Perfect, Innovate

Measurement System Analysis (MSA)

(a.k.a. Gage Capability Analysis)

Total (recorded) Variability is broken into two major pieces:



Product / Service Variability | **Measurement Variability**

- **MSA identifies and quantifies the different sources of variation that affect a measurement system.**
- **Variation in measurements can be attributed to variation in the product/service itself or to variation in the measurement system.**
- **The variation in the measurement system itself is measurement error.**

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Measurement Variability Broken Down Further

Purpose:

To assess how much variation is associated with the measurement system and to compare it to the total process variation or tolerances.

$$\sigma^2_{total} = \sigma^2_{product} + \sigma^2_{measurement}$$

$$\sigma^2_{measurement} = \sigma^2_{repeatability} + \sigma^2_{reproducibility}$$

Repeatability:

Variation obtained by the same person using the same procedure on the same product, transaction or service for repeated measurements (variability ***within*** operator).

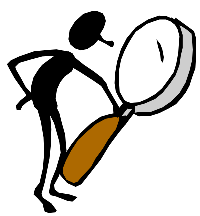
Reproducibility:

Variation obtained due to differences in people who are taking the measurements (variability ***between*** operators).

More About Reproducibility and Repeatability

For the scenario below, look at the data and indicate which of the following you think is true:

a. Repeatability appears to be more of a problem than reproducibility
b. Reproducibility appears to be more of a problem than repeatability
c. Repeatability and Reproducibility appear to be about the same



	Operator 1		Operator 2	
	Rep 1	Rep 2	Rep 1	Rep 2
Part 1	21	23	26	28
Part 2	19	18	24	24
Part 3	20	23	27	24
Part 4	19	22	21	20

- MSA will help quantify, more exactly, the capability of the measurement system and answer questions about repeatability, reproducibility, and capability with respect to the customer specs.

- **Rule of Thumb for MSA Sample Size:**
  - Variables (Continuous) Data: (# operators)*(number of parts) ≥ 20
  - Attribute (Binary) Data: (# operators)*(number of parts) ≥ 60

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Commonly Used Measures for Diagnosing Measurement Systems

1. Precision-to-Tolerance Ratio (P/TOL)

$$P/TOL = \frac{6\sigma_{meas}}{USL - LSL}$$ (Specification Limits are needed)

ROT: If P/TOL $\leq$ .10 : Very Good Measurement System

P/TOL $\geq$ .30 : Unacceptable Measurement System

2. Precision-to-Total Ratio (P/TOT)

$$P/TOT = \frac{\sigma_{meas}}{\sigma_{total}}$$

ROT: If P/TOT $\leq$ .10 : Very Good Measurement System

P/TOT $\geq$ .30 : Unacceptable Measurement System

3. Discrimination or Resolution $= \left(\frac{\sigma_{product}}{\sigma_{meas}}\right) \times 1.41$

(# of truly distinct measurements that can be obtained by the measurement system) ROT: Resolution ≥ 5

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Analysis of Variance (ANOVA) Results

(from SPC XL)

MSA ANOVA Method Results

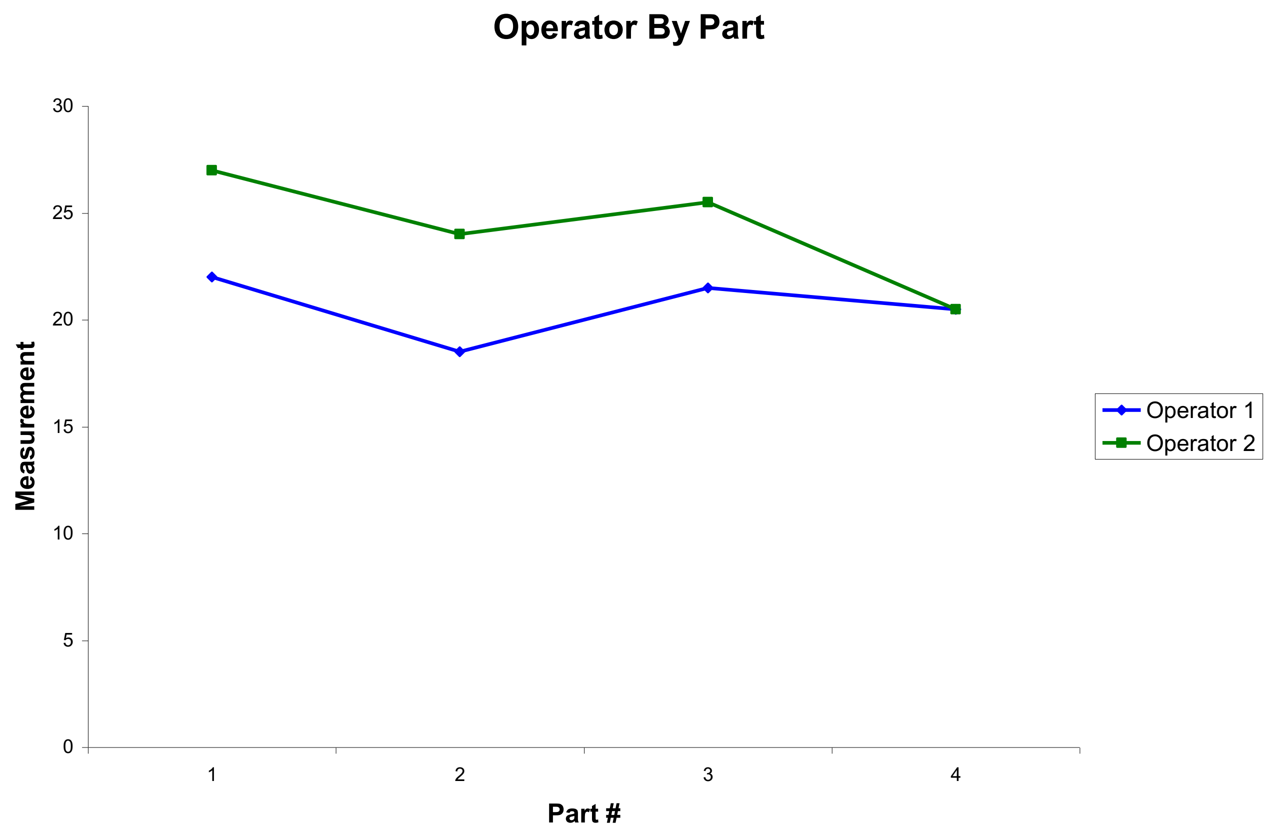
Source	Variance	Standard Deviation	% Contribution	p Value
Total Measurement (Gage)	10.0625	3.172144385	83.71%	
Repeatability	2.3125	1.520690633	19.24%	
Reproducibility	7.75	2.783882181	64.47%	
Operator	5.79166667	2.40658818	48.18%	
Oper * Part Interaction	1.95833333	1.399404635	16.29%	0.1167
Product (Part-to-Part)	1.95833333	1.399404635	16.29%	
Total	12.0208333	3.467107344	100.00%	

USL	35
LSL	10
Precision to Tolerance Ratio	0.76131465
Precision to Total Ratio	0.91492535
Resolution	0.6

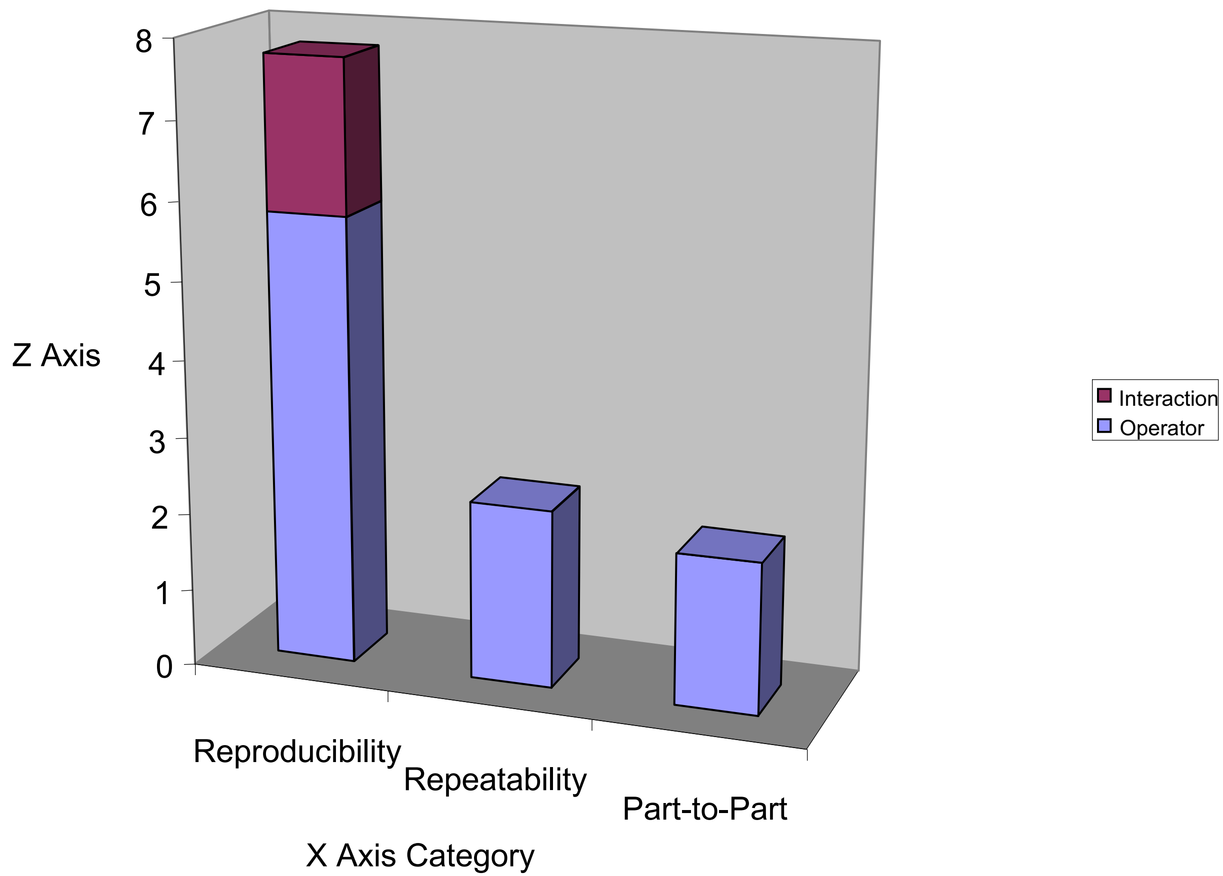
BIAS ANALYSIS	
Reference	Bias
Not Available	

- P/TOL and P/TOT are too high.
- Resolution is unacceptable.
- Reproducibility is significantly larger than Repeatability and appears to be the biggest problem with this measurement process.

Graphical View of Operator by Part Interaction
(from SPC XL)



Graphical View of Variance Components
(from SPC XL)



Measurement System Analysis Exercise #2

Statapult® Catapult

Data Collection Template #2

Shot #	Operator 1	Operator 2	Operator 3
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Air Academy Associates
Simplify, Perfect, Innovate

Motivation for DOE from Dr. Gilmore (DOT&E)

(from his 26 June 2013 memo on Flawed Applications of DOE)

1. One of the most important goals of operational testing is to **characterize** a system's effectiveness over the operational envelope.

2. I advocate the use of DOE to ensure that test programs are able to **determine the effect of factors on** a comprehensive set of operational mission-focused and **quantitative response variables.**

3. Future test plans must state clearly that data are being collected to measure a particular response variable (possibly more than one) in order to **characterize** the system's performance by examining the effects of multiple factors … and clearly delineating what statistical **model** (e.g., main effects and interactions) is motivating … the variation of the test.

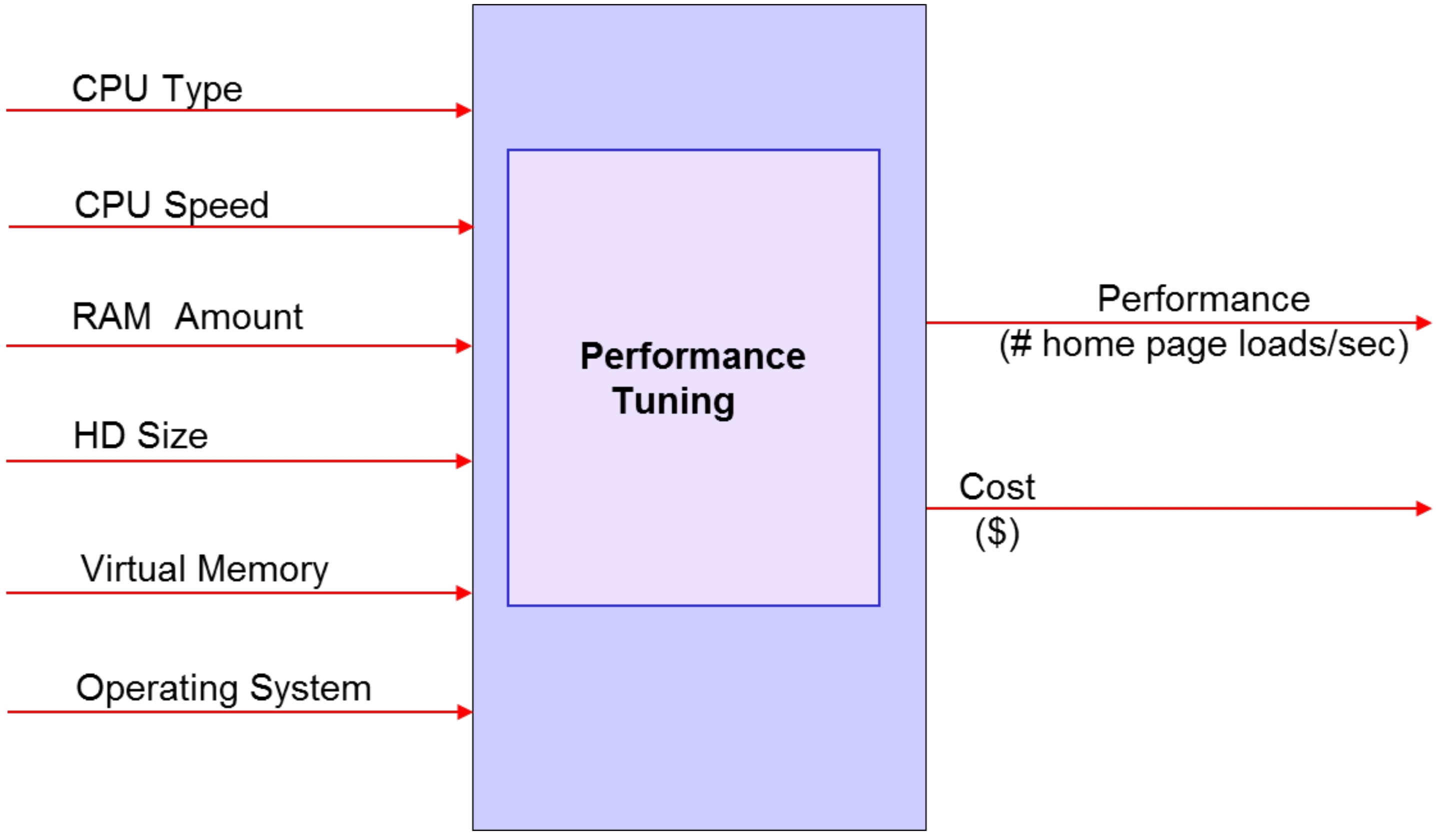
4. Confounding factors must be avoided.

5. Another pitfall to avoid is relying on binary metrics as the primary response variable.

Air Academy Associates

Simplify, Perfect, Innovate

Web-Based Test Scenario



Simplify, Perfect, Innovate

Combinatorial Test/DOE Terminology

Y: Output, response variable, dependent variable

X: Input, factor, independent variable (a measurable entity that is purposely changed during an experiment)

Level: A unique value or choice of a factor (X)

Run: An experimental combination of the levels of the X's

Replication: Doing or repeating an experimental combination

Effect: The difference or impact on Y when changing X

Interaction: When the effect of one factor depends on the level of another factor

Air Academy Associates
Simplify, Perfect, Innovate

Performance Tuning Terminology

Factors/Inputs (X's)	Levels (Choices)	Performance/Outputs (Y's)
CPU Type	Itanium, Xeon	# home page loads/sec
CPU Speed	1 GHz, 2.5 GHz	Cost
RAM Amount	256 MB, 1.5 GB	
HD Size	50 GB, 500 GB	
VM	J2EE, .NET	
OS	Windows, Linux	

Which factors are important? Which are not?
Which combination of factor choices will create operational problems?
How do you know for sure? Show me the data.

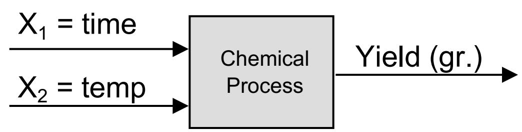
Air Academy Associates
Simplify, Perfect, Innovate

Approaches to Testing Multiple Factors

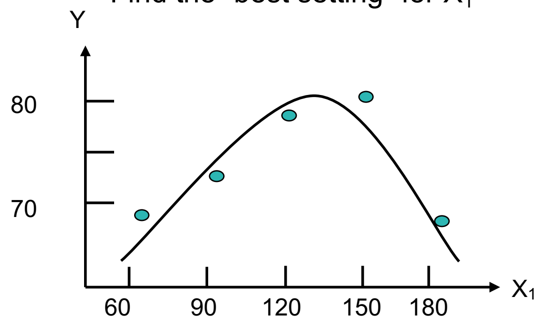
- **Traditional Approaches**
  - One Factor at a Time (OFAT)
  - Oracle (Best Guess)
  - All possible combinations (full factorial)
- **Modern Approach**
  - Statistically designed experiments (DOE) … full factorial plus other selected DOE designs, depending on the situation

Air Academy Associates
Simplify, Perfect, Innovate

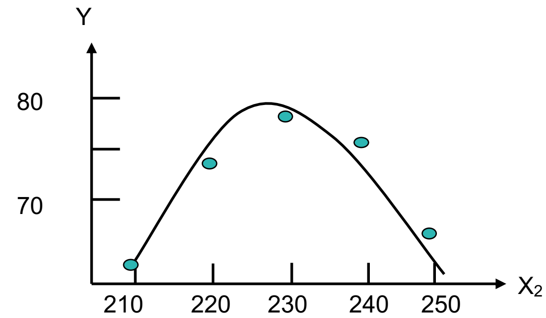
OFAT (One Factor at a Time)



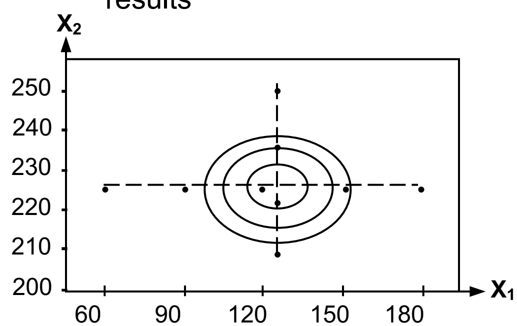
1. Hold X_2 constant and vary X_1
Find the "best setting" for X_1



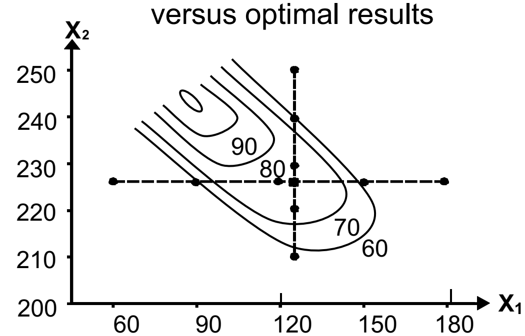
2. Hold X_1 constant at "best setting" and vary X_2.
Find the "best setting" for X_2.



3. One factor at a time results



4. One factor at a time results versus optimal results



The Good and Bad about OFAT

- **Good News**
  - Simple
  - Intuitive
  - The way we were originally taught
- **Bad News**
  - Will not be able estimate variable interaction effects
  - Will not be able to generate prediction models and thus not be able to optimize performance

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Oracle (Best Guess)

X1 = W = Wetting Agent (1=.07 ml; 2=none)

X2 = P = Plasticizer (1=1ml; 2=none)

X3 = E = Environment (1=Ambient Mixing; 2=Semi-Evacuated)

X4 = C = Cement (1=Portland Type III; 2=Calcium Aluminate)

X5 = A = Additive (1=No Reinforcement; 2=Steel)

Y = Strength of Lunar Concrete

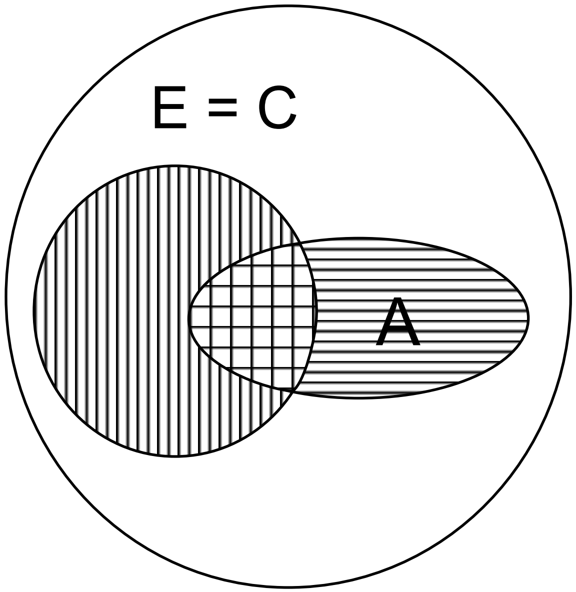
Run	W	P	E	C	A	Y
1	1	2	1	1	1	5
2	1	1	1	1	1	6
3	2	2	1	1	1	5
4	2	1	1	1	2	6
5	1	2	2	2	2	7
6	1	1	2	2	2	8
7	2	2	2	2	2	10
8	2	1	2	2	1	11

Does factor C shift the average of Y?

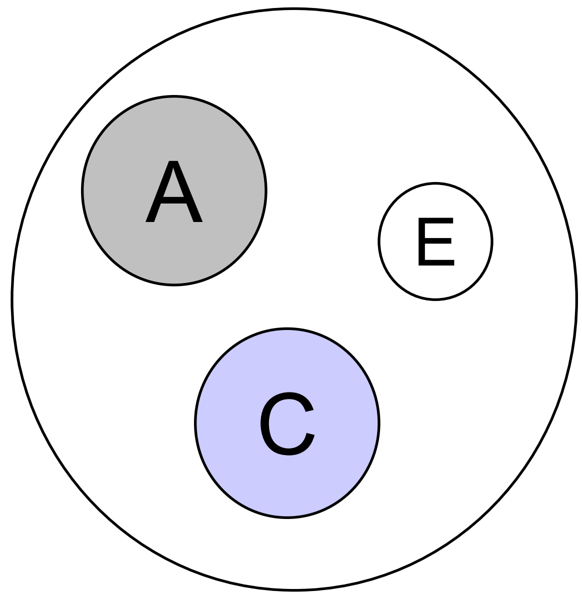
Air Academy Associates
Simplify, Perfect, Innovate

Evaluating the Effects of Variables on Y

What we have is:



What we need is a design to provide independent estimates of effects:



How do we obtain this independence of variables?

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Statistically Designed Experiments (DOE):
Orthogonal or Nearly Orthogonal Designs

- FULL FACTORIALS (for small numbers of factors)
- FRACTIONAL FACTORIALS
- PLACKETT - BURMAN
- LATIN SQUARES
- HADAMARD MATRICES

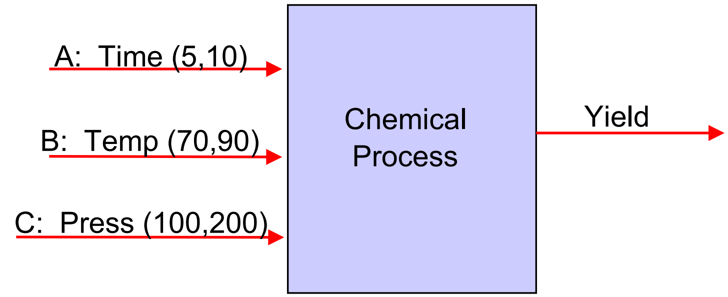
(Plackett - Burman, Latin Squares, Hadamard Matrices: Taguchi Designs)

- BOX - BEHNKEN DESIGNS
- CENTRAL COMPOSITE DESIGNS

(Box - Behnken, Central Composite: Response Surface Designs)

- HIGH THROUGHPUT TESTING (ALL PAIRS)
- NEARLY ORTHOGONAL LATIN HYPERCUBE DESIGNS

Example:



SIMPLE DEFINITION OF A TWO-LEVEL ORTHOGONAL DESIGN

Run	Actual Settings (5, 10) A: Time	Actual Settings (70, 90) B: Temp	Actual Settings (100, 200) C: Press	Coded Matrix (A) Time	Coded Matrix (B) Temp	Coded Matrix (C) Press	Interactions (AB) Uncoded	Interactions (AB) Coded
1	5	70	100	-1	-1	-1	350	+1
2	5	70	200	-1	-1	+1	350	+1
3	5	90	100	-1	+1	-1	450	-1
4	5	90	200	-1	+1	+1	450	-1
5	10	70	100	+1	-1	-1	700	-1
6	10	70	200	+1	-1	+1	700	-1
7	10	90	100	+1	+1	-1	900	+1
8	10	90	200	+1	+1	+1	900	+1

The Beauty of Orthogonality
(Vertical and Horizontal Balance)

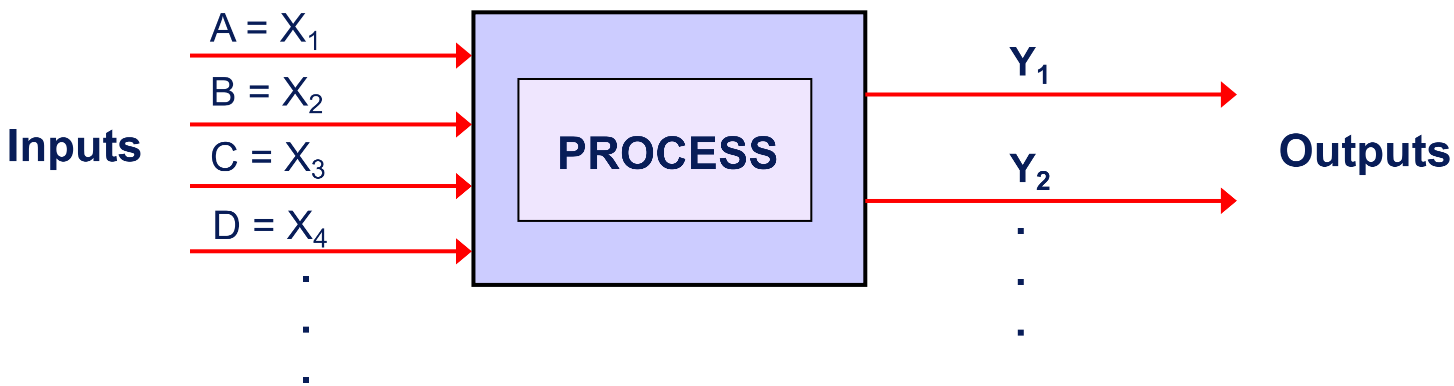
A Full Factorial Design for 3 Factors A, B, and C, Each at 2 levels:

Run	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
2	-	-	+	+	-	-	+
3	-	+	-	-	+	-	+
4	-	+	+	-	-	+	-
5	+	-	-	-	-	+	+
6	+	-	+	-	+	-	-
7	+	+	-	+	-	-	-
8	+	+	+	+	+	+	+

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

What is a Designed Experiment?

Purposeful changes of the inputs (factors) in order to observe corresponding changes in the output (response).



Run	X_1	X_2	X_3	X_4	Y_1	Y_2	. . . . . .	$\bar{Y}$	S_Y
1									
2									
3									
.									
.									

AIR ACADEMY ASSOCIATES

Simplify, Perfect, Innovate

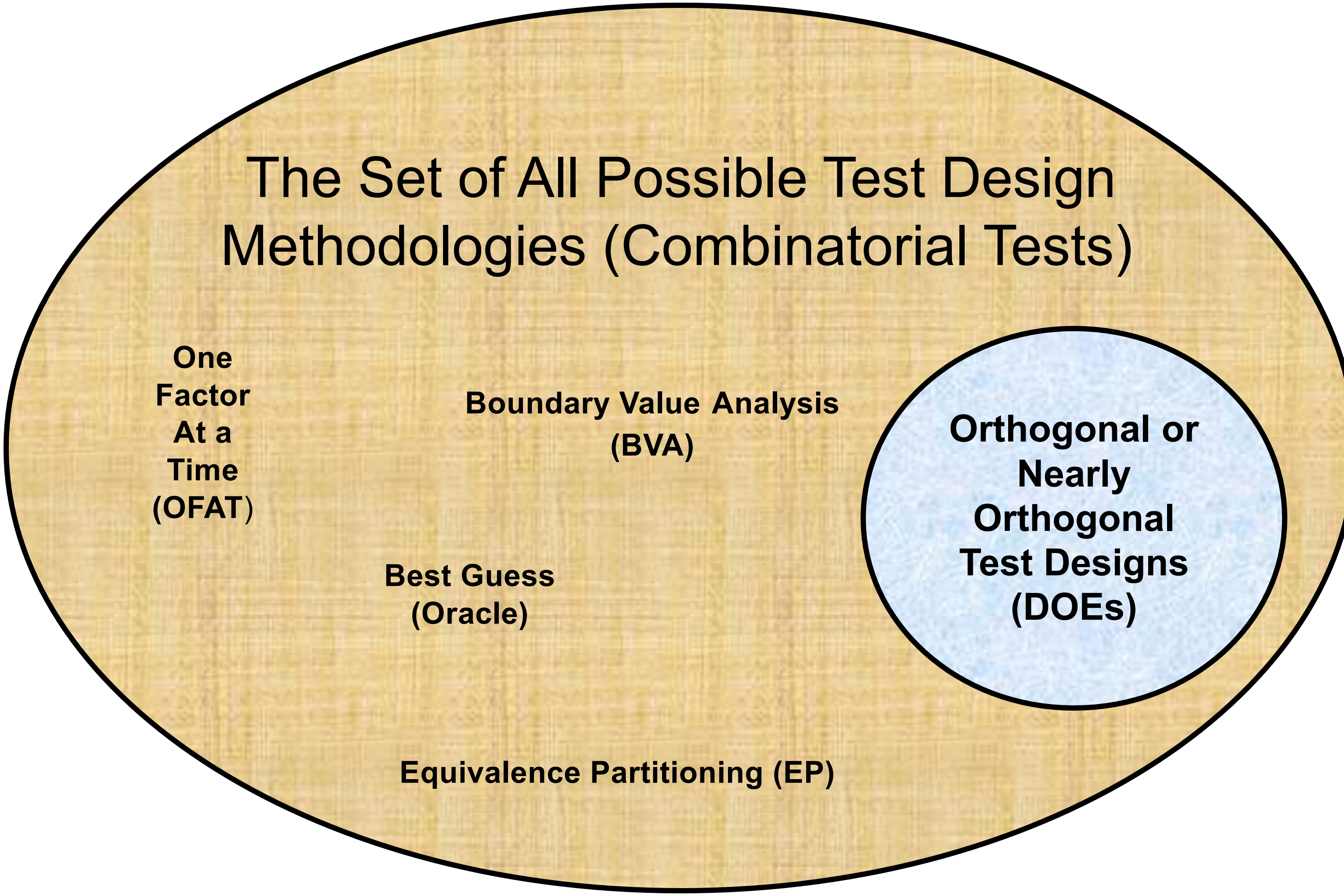
Famous Quote

"All experiments are designed experiments; some are poorly designed, some are well designed."

George Box (1919-2013), Professor of Statistics, DOE Guru

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Design of Experiments (DOEs): A Subset of All Possible Test Design Methodologies



AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

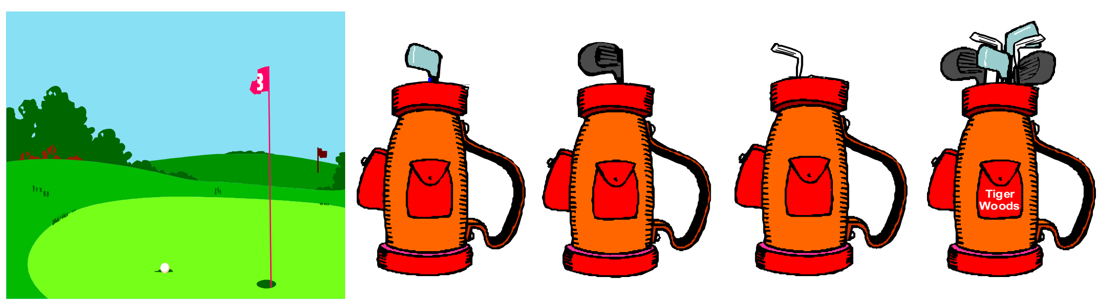
Design of Experiments (DOE)

- An optimal data collection methodology
- “Interrogates” the process
- Used to identify important relationships between input and output factors
- Identifies important interactions between process variables
- Can be used to optimize a process
- Changes “I think” to “I know”

Air Academy Associates
Simplify, Perfect, Innovate

Important Contributions From:

	TAGUCHI	SHAININ	CLASSICAL	BLENDED APPROACH
Loss Function	*			*
Emphasis on Variance Reduction	*			*
Robust Designs	*			*
KISS	*	*		*
Simple Significance Tests		*		*
Component Swapping		*		*
Multivariate Charts		*		*
Modeling			*	*
Sample Size			*	*
Efficient Designs			*	*
Optimization			*	*
Confirmation	*			*
Response Surface Methods			*	*



Which bag would a world class golfer prefer?

AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

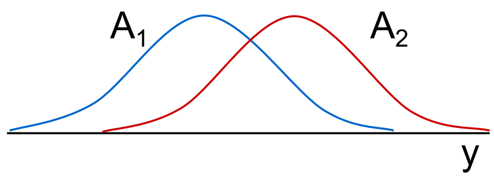
Three Major Reasons for Using a DOE

- **Screening**
  - For testing many factors in order to **separate** the critical factors from the trivial many.
- **Modeling**
  - For building **functions** that can be used to predict outcomes, assess risk, and optimize performance. These **include** the ability to evaluate **interaction and higher order effects**.
- **Performance Verification and Validation**
  - For **confirming** that a system performs in accordance with its specifications/requirements.

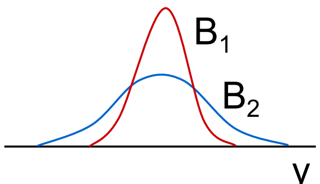
Air Academy Associates
Simplify, Perfect, Innovate

DOE Helps Determine How Inputs Affect Outputs

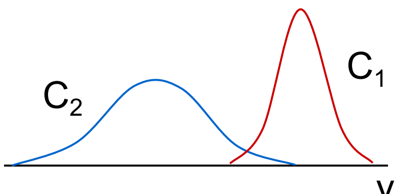
i) Factor A affects the average of y



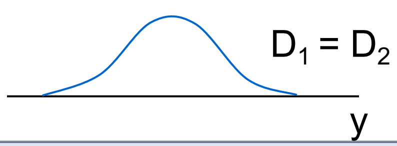
ii) Factor B affects the standard deviation of y



iii) Factor C affects the average and the standard deviation of y

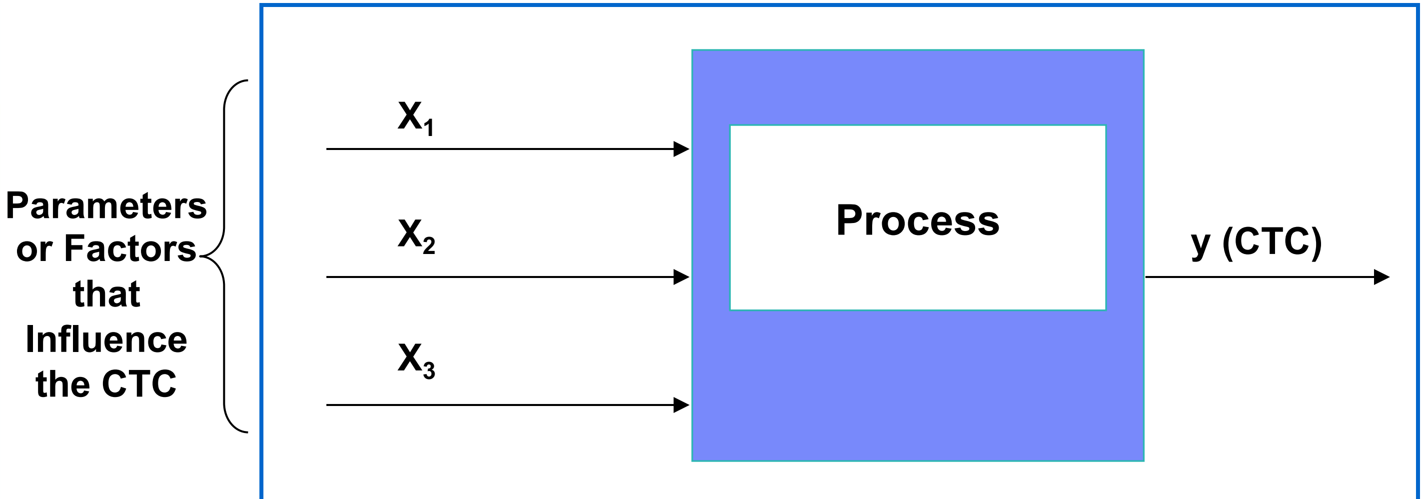


iv) Factor D has no effect on y



AIR ACADEMY ASSOCIATES
Simplify, Perfect, Innovate

Transfer Functions



$$\hat{y} = f_1 (x_1, x_2, x_3)$$

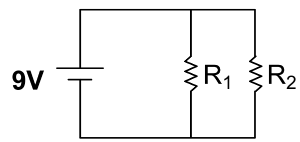
$$\hat{s} = f_2 (x_1, x_2, x_3)$$

Where does the transfer function come from?

- **Exact transfer Function**
- **Approximations**
  - **DOE**
  - **Historical Data Analysis**
  - **Simulation**

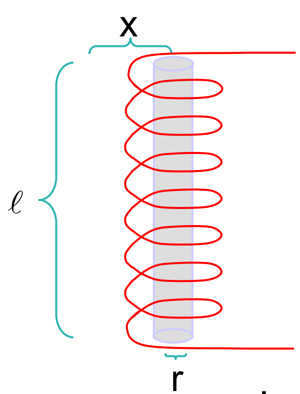
Exact Transfer Functions

- Engineering Relationships
  - V = IR
  - F = ma



The equation for current (I) through this DC circuit is defined by:

$$I = \frac{V}{\frac{R_1 \cdot R_2}{R_1 + R_2}} = \frac{V(R_1 + R_2)}{R_1 \cdot R_2}$$



The equation for magnetic force at a distance X from the center of a solenoid is:

$$H = \frac{NI}{2\ell}\left[\frac{.5\ell + x}{\sqrt{r^2 + (.5\ell + x)^2}} + \frac{.5\ell - x}{\sqrt{r^2 + (.5\ell - x)^2}}\right]$$

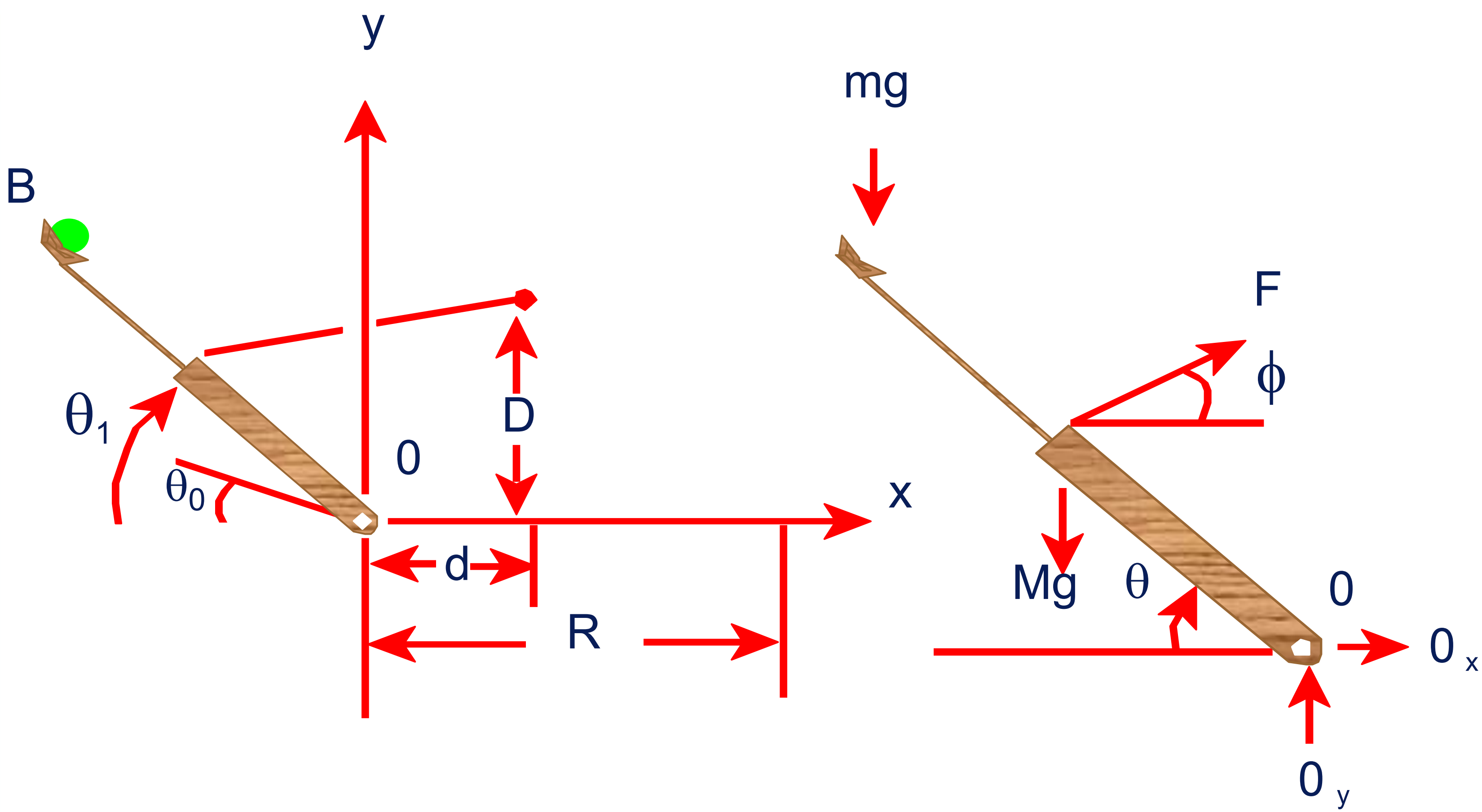
where

N : total number of turns of wire in the solenoid
I : current in the wire, in amperes
r : radius of helix (solenoid), in cm
ℓ : length of the helix (solenoid), in cm
x : distance from center of helix (solenoid), in cm
H : magnetizing force, in amperes per centimeter

Catapulting Power into Test and Evaluation

Statapult® Catapult

The Theoretical Approach



The Theoretical Approach (cont.)

$$I_0\ddot{\theta} = r_F F(\theta)\sin\theta\cos\varphi - (Mgr_G + mgr_B)\sin\theta \qquad \tan\phi = \frac{D - r_F\sin\theta}{d + r_F\cos\theta},$$

$$\frac{1}{2}I_0\dot{\theta}^2 = r_F\int_{\theta_0}^{\theta} F(\theta)\sin\theta\cos\varphi\, d\theta - (Mgr_G + mgr_B)(\sin\theta - \sin\theta_0)$$

$$\frac{1}{2}I_0\dot{\theta}_1^2 = r_F\int_{\theta_0}^{\theta_1} F(\theta)\sin\theta\cos\varphi\, d\theta - (Mgr_G + mgr_B)(\sin\theta_1 - \sin\theta_0).$$

$$x = v_B\cos\left(\frac{\pi}{2} - \theta_1\right)t - \frac{1}{2}r_B\cos\theta_1 \qquad y = r_B\sin\theta_1 + v_B\sin\left(\frac{\pi}{2} - \theta_1\right)t - \frac{1}{2}gt^2.$$

$$r_B\sin\theta_1 + (R + r_B\cos\theta_1)\tan\left(\frac{\pi}{2} - \theta_1\right) - \frac{g}{2V_B^2}\frac{(R + r_B\cos\theta_1)^2}{\cos^2\left(\frac{\pi}{2} - \theta_1\right)} = 0.$$

$$\frac{gI_0}{4r_B}\frac{(R + r_B\cos\theta_1)^2}{\cos^2\left(\frac{\pi}{2} - \theta_1\right)\left[r_B\sin\theta_1 + (R + r_B\cos\theta_1)\tan\left(\frac{\pi}{2} - \theta_1\right)\right]}$$

$$= r_F\int_{\theta_0}^{\theta_1} F(\theta)\sin\theta\cos\phi\, d\theta - (Mgr_G + mgr_B)(\sin\theta_1 - \sin\theta_0).$$

Statapult® DOE Demo
(The Empirical Approach)

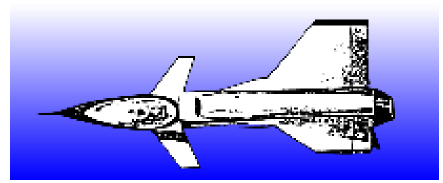
	Actual Factors		Coded Factors			Response Values			
Run	A	B	A	B	AB	Y_1	Y_2	$\bar{Y}$	S
1	144	2	-1	-1	+1				
2	144	3	-1	+1	-1				
3	160	2	+1	-1	-1				
4	160	3	+1	+1	+1				

Simplified Table for Determining Sample Size Based on Confidence and Power

Percent Confidence that a term identified as significant, truly does belong in $\hat{s}$ [$\hat{y}$]	Percent chance of finding a significant variance [average] shifting term if one actually exists	Number of Runs in 2 Level Portion of the Design				
		2	4	8	12	16
		Sample Size per Experimental Condition				
95% (α = .05)	40% (β = .60)	5 [3]	3 [2]	2 [1]	N/A	N/A
95% (α = .05)	75% (β = .25)	9 [5]	5 [3]	3 [2]	2 [1]	2 [1]
95% (α = .05)	90% (β = .10)	13 [7]	7 [4]	4 [2]	3 [2]	N/A
95% (α = .05)	95% (β = .05)	17 [9]	9 [5]	5 [3]	4* [2]	3 [2]
95% (α = .05)	99% (β = .01)	21 [11]	11 [6]	6 [3]	5* [3]	4* [2]

Air Academy Associates
Simplify, Perfect, Innovate

Value Delivery: Reducing Time to Market for New Technologies



INPUT

OUTPUT

Pitch <) (0, 15, 30)

Roll <) (0, 15, 30)

W1F <) (-15, 0, 15)

W2F <) (-15, 0, 15)

W3F <) (-15, 0, 15)

Modeling Flight Characteristics of New 3-Wing Aircraft

Six Aero-Characteristics

- **Total # of Combinations = 3^5 = 243**
- **Central Composite Design: n = 30**

Patent Holder: Dr. Bert Silich

Air Academy Associates

Simplify, Perfect, Innovate

Aircraft Equations

C_L = .233 + .008(P)^2 + .255(P) + .012(R) - .043(WD1) - .117(WD2) + .185(WD3) + .010(P)(WD3) - .042(R)(WD1) + .035(R)(WD2) + .016(R)(WD3) + .010(P)(R) - .003(WD1)(WD2) - .006(WD1)(WD3)

C_D = .058 + .016(P)^2 + .028(P) - .004(WD1) - .013(WD2) + .013(WD3) + .002(P)(R) - .004(P)(WD1) - .009(P)(WD2) + .016(P)(WD3) - .004(R)(WD1) + .003(R)(WD2) + .020(WD1)^2 + .017(WD2)^2 + .021(WD3)^2

C_Y = -.006(P) - .006(R) + .169(WD1) - .121(WD2) - .063(WD3) - .004(P)(R) + .008(P)(WD1) - .006(P)(WD2) - .008(P)(WD3) - .012(R)(WD1) - .029(R)(WD2) + .048(R)(WD3) - .008(WD1)^2

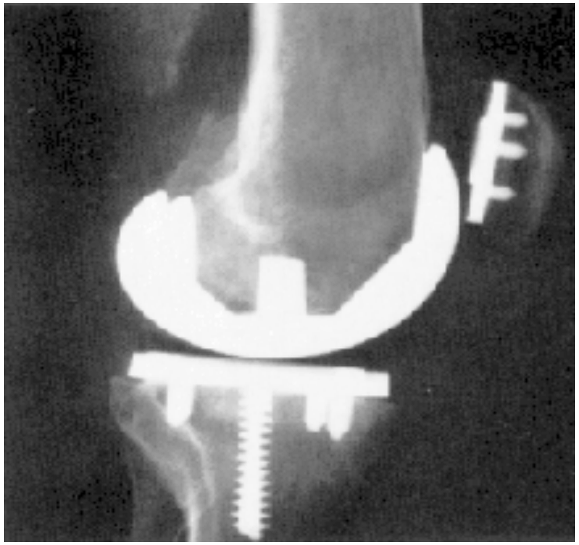
C_M = .023 - .008(P)^2 + .004(P) - .007(R) + .024(WD1) + .066(WD2) - .099(WD3) - .006(P)(R) + .002(P)(WD2) - .005(P)(WD3) + .023(R)(WD1) - .019(R)(WD2) - .007(R)(WD3) + .007(WD1)^2 - .008(WD2)^2 + .002(WD1)(WD2) + .002(WD1)(WD3)

C_{YM}= .001(P) + .001(R) - .050(WD1) + .029(WD2) + .012(WD3) + .001(P)(R) - .005(P)(WD1) - .004(P)(WD2) - .004(P)(WD3) + .003(R)(WD1) + .008(R)(WD2) - .013(R)(WD3) + .004(WD1)^2 + .003(WD2)^2 - .005(WD3)^2

C_e = .003(P) + .035(WD1) + .048(WD2) + .051(WD3) - .003(R)(WD3) + .003(P)(R) - .005(P)(WD1) + .005(P)(WD2) + .006(P)(WD3) + .002(R)(WD1)

Air Academy Associates
Simplify, Perfect, Innovate

Fusing Titanium and Cobalt-Chrome



Courtesy Rai Chowdhary

Google on DOE

(quotes* from Daryl Pregibon, Google Engineer)

"From a user's perspective, a query was submitted and results appear. From Google's perspective, the user has provided an opportunity to test something. What can we test? Well, there is so much to test that we have an Experiment Council that vets experiment proposals and quickly approves those that pass muster."

" We evangelize experimentation to the extent that we provide a mechanism for advertisers to run their own experiments.

. . . allows an advertiser to run a (full) factorial experiment on its web page. Advertisers can explore layout and content alternatives while Google randomly directs queries to the resulting treatment combinations. Simple analysis of click and conversion rates allows advertisers to explore a range of alternatives and their effect on user awareness and interest."

*** Taken From: *Statistics @ Google* in Amstat News, May 2011**

Air Academy Associates

Simplify, Perfect, Innovate

Best Practices for “Operationalizing” DOE

(i.e., changing the culture to one of habitually using DOE)

1. **Coaching on projects is an absolute must.**

2. **A Keep-It-Simple-Statistically (KISS) approach with easy-to-comprehend materials and easy-to-use software.**

3. **Gaining and propagating quick-hitting successes.**

4. **Getting leadership on board and continuously re-invigorating them is necessary.**

5. **Developing a culture of continuously generating transfer functions for the purpose of optimization, prediction, and risk assessment.**

Air Academy Associates

Simplify, Perfect, Innovate

Key Take-Aways

- STAT includes various techniques such as DOE, MSA, and hypothesis testing.
- Hypothesis testing allows us to control, via sample size, the P(false detection) and P(missed detection), the Type I (alpha risk) and Type II (beta risk) errors, respectively.
- Measurement System Analysis (MSA) is a test on the measurement system itself.
- MSA will help quantify, more exactly, the capability of the measurement system and answer questions about repeatability, reproducibility, and capability with respect to the customer specs.
- There are various approaches to testing many factors simultaneously, to include OFAT and Oracle (Best Guess).
- DOE brings orthogonal or nearly orthogonal designs into play and can be used to screen, model, or perform validation testing.
- DOE is the key link between Test and Evaluation, because it allows us to evaluate the effects of factors and their interactions independently from one another.
- Learning about DOE and making it practical in an organization does NOT have to be difficult. Following the 5 Best Practices for "operationalizing" DOE in an organization can make it happen.

Air Academy Associates

Simplify, Perfect, Innovate

CAUTION!!

This presentation is designed to be an overview of the topic of STAT. It is not meant to make anyone an expert in the subject matter. Becoming a practitioner will require more education than what this tutorial provides.

Air Academy Associates will be offering the week-long Scientific Test and Analysis Techniques (STAT) course in Colorado Springs the weeks of 22-26 June and 19-23 October, 2015. Your participation in this tutorial earns you a significant reduction in price to this class.

For more information, please contact

Air Academy Associates, LLC
1650 Telstar Drive, Ste 110
Colorado Springs, CO 80920

Toll Free: (800) 748-1277 or (719) 531-0777
Facsimile: (719) 531-0778
Email: aaa@airacad.com
or mkiemele@airacad.com
Website: www.airacad.com

Air Academy Associates
Simplify, Perfect, Innovate