Application of Simulation to Sample Size Calculations and Teaching Statistical Concepts

Fall Technical Conference

Statistics: Powering a Revolution in Quality Improvement

October 4 - 6, 2017 | Philadelphia, PA

Presented by:
Lou Johnson
lou@jmcdataexperts.com

JMC Data Experts

Question: Which Door Should I Take?

<u>One Car and Two Goats</u>

1. Contestant choses a door.
2. Monty opens one door with a goat.
3. Monty offers the contestant the remaining door or stay with their door.

1. Switch to the other door ?

2. Stay with your door?

3. It doesn’t matter?

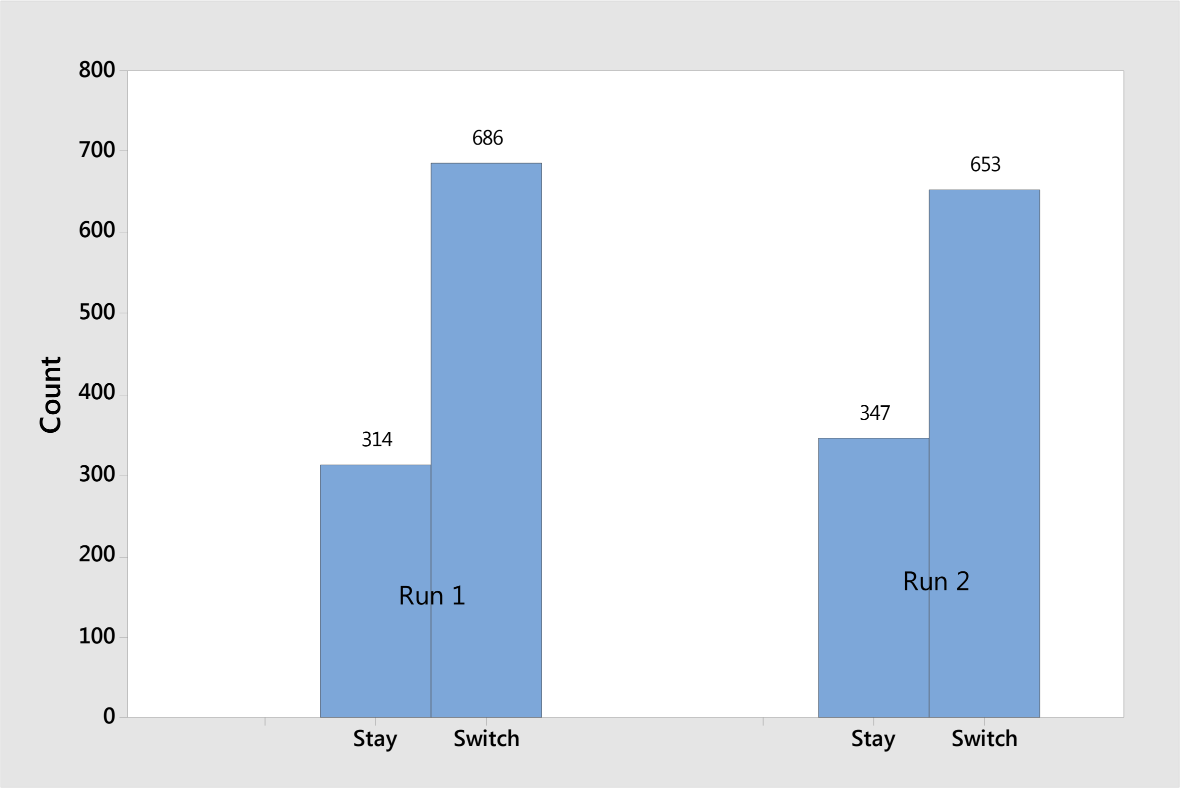
The Right Answer Explained (?)

Since you seem to enjoy coming straight to the point, I'll do the same. You blew it! Let me explain. If one door is shown to be a loser, that information changes the probability of either remaining choice, neither of which has any reason to be more likely, to 1/2. As a professional mathematician, I'm very concerned with the general public's lack of mathematical skills. Please help by confessing your error and in the future being more careful.
Robert Sachs, Ph.D.
George Mason University

You blew it, and you blew it big! Since you seem to have difficulty grasping the basic principle at work here, I'll explain. After the host reveals a goat, you now have a one-in-two chance of being correct. Whether you change your selection or not, the odds are the same. There is enough mathematical illiteracy in this country, and we don't need the world's highest IQ propagating more. Shame!
Scott Smith, Ph.D.
University of Florida

Could a Quick Simulation Answer the Question Definitively?

```
GMacro
Goat3
Name C1 "Door"
Name C2 "Car"
Name c3 "Guess"
Name c4 "Wins"
Do k1 = 1 : 1000
Sample 1 'Door' 'Car'
Sample 1 'Door'  'Guess'
Let c4[k1] = 'Car' = 'Guess'
Enddo
Name k2 "Stay Wins"
Name k3 "Switch Wins"
Let k2 = Sum(c4)
Let k3 = 1000 - k2
Print k2 k3
endmacro
```



https://math.ucsd.edu/~crypto/Monty/montybg.html

Agenda

- The Monty Hall Paradox
- Confidence Intervals for the Mean and Standard Deviation
  - Percent variability for the mean and standard deviation
  - Impact on process capability
  - Effect on MSA measures of capability and measuring repeatability
- Sampling Plans for Expanded Gage R & R
  - Simulation results for sources of variation
  - The best sampling plan when variation sources cannot be estimated
- Analysis of Variance Study when Factors are Nested
  - Syringe needle strength – where is the variability coming from?
  - Study results for the complete study
  - Simulation results on the effect of sample size on study results
- Cleaning Process Factorial with a Non-normal Covariate
  - Factorial design for a cleaning process with a non-normal covariate
  - Power and sample size calculations
  - Simulation for the effect of sample size on averaging the covariate
- Questions & Discussion

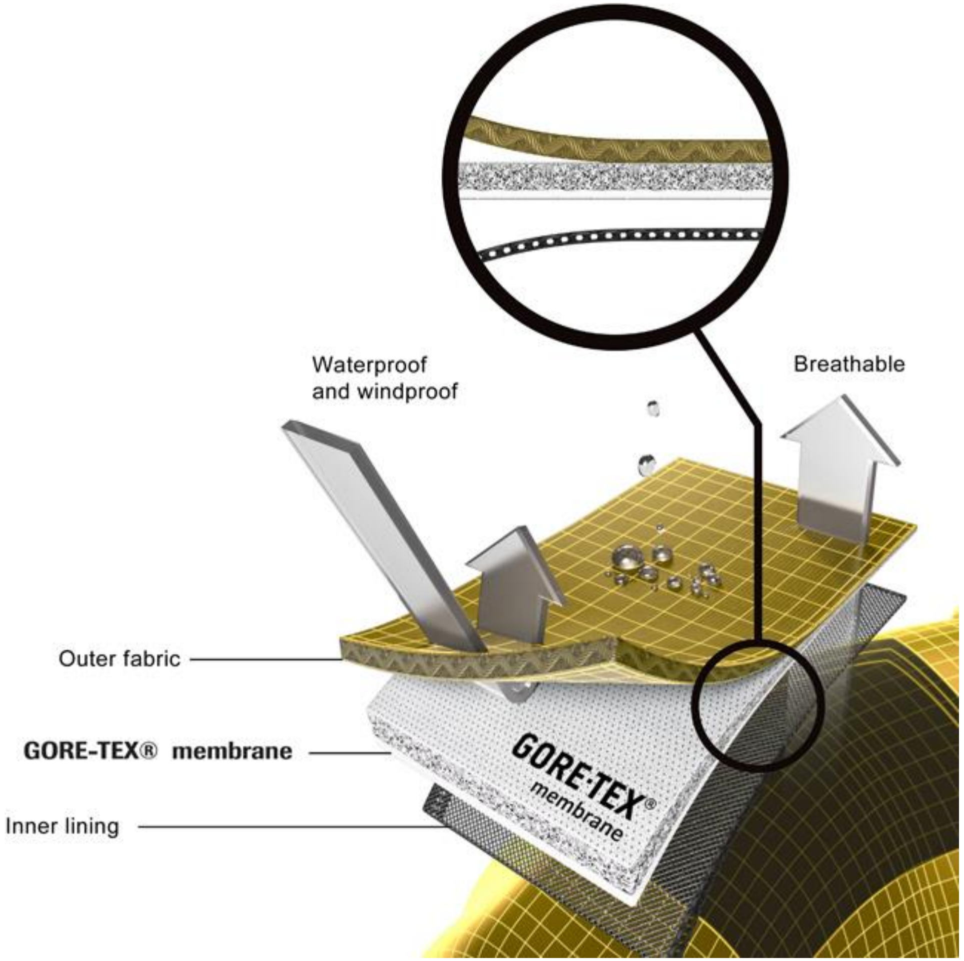
Non-disclosure Statement

"YOU CAN TELL ME ANYTHING, MRS. ROJAK. I'D NEVER VIOLATE THE SACRED FISH STORE-CUSTOMER RELATIONSHIP."

All industrial experiments, results and scenarios are based on the authors' actual experiences.
Data units, variable names, etc have been changed for demonstration purposes only.

JMC Data Experts

The Difference between a Statistician and an Engineer



The engineer takes three measurements (5.5, 6.1, 6.4) , takes the average and concludes that the thickness of the fabric is 6.0 /1000", give or take.

The statistician concludes that the engineer took three samples from a population that is likely normally distributed, centered at 6.0 /1000" and has a standard deviation of 1 /1000".

The difference is critical. The second interpretation understands that future samples from the population will be different even if nothing has changed.

How do We Teach this Important Concept of Variability

Minitab - Simulation for Mean and Standard Deviation.mpj -

File Edit Data Calc Stat Graph Editor Tools W

↓	C1	C2	C3	C4	C5
	Population	Sample			
1	7.44434	6.06817			
2	7.27192	5.58695			
3	4.34465	6.19379			
4	7.46776	6.43897			
5	4.78513	6.29044			
6	5.95534	7.77908			
7	5.09990	5.95274			
8	6.92940	5.93298			
9	4.88650				
10	4.03013				
11	7.49095				
12	4.42031				
13	6.22871				
14	6.04138				
15	5.93097				
16	5.28613				
17	6.32382				
18	6.18136				

Classroom Exercise

1. Create a population with a given mean and std deviation.
2. Take a random sample of 8 measurements.
3. Calculate the descriptive statistics.

Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Sample	8	5.169	1.653	2.242	7.128

Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Sample	8	6.280	0.658	5.587	7.779

Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Sample	8	5.923	1.051	4.935	7.542

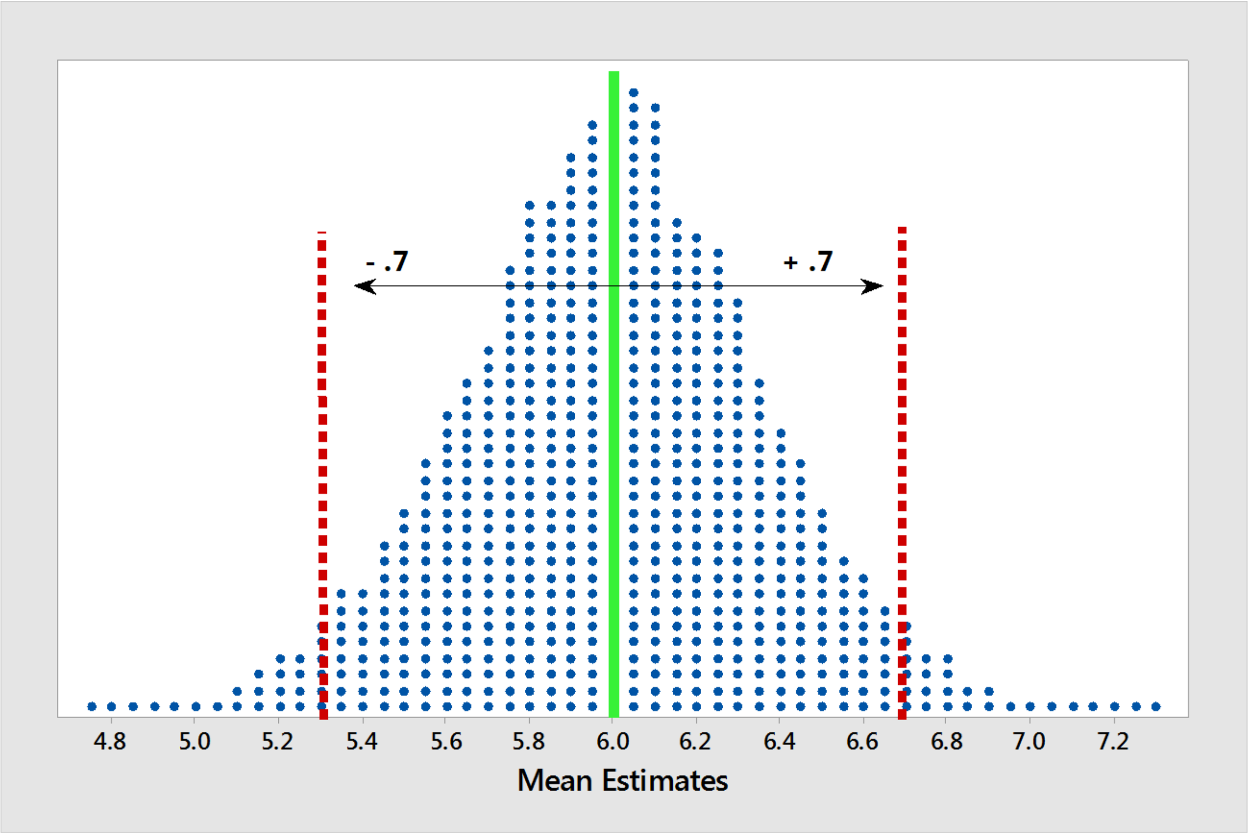
Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Sample	8	6.382	0.653	5.110	7.013

Repeat sampling 10,000 times. What is the distribution of the estimates of the mean?

JMC Data Experts

Distribution of 10,000 Estimates of the Mean



Conclusions

1. 95% of the estimates of the mean fall between 5.3 and 6.7.
2. Your estimate from 8 samples will likely be within + / - 12% of the right answer.
3. The confidence interval for the mean is

   Lower limit = M - $Z_{.95}\sigma_M$

   Upper limit = M + $Z_{.95}\sigma_M$

(6 – 1.96 * .3536 , 6 , 6 + 1.96*.3536)

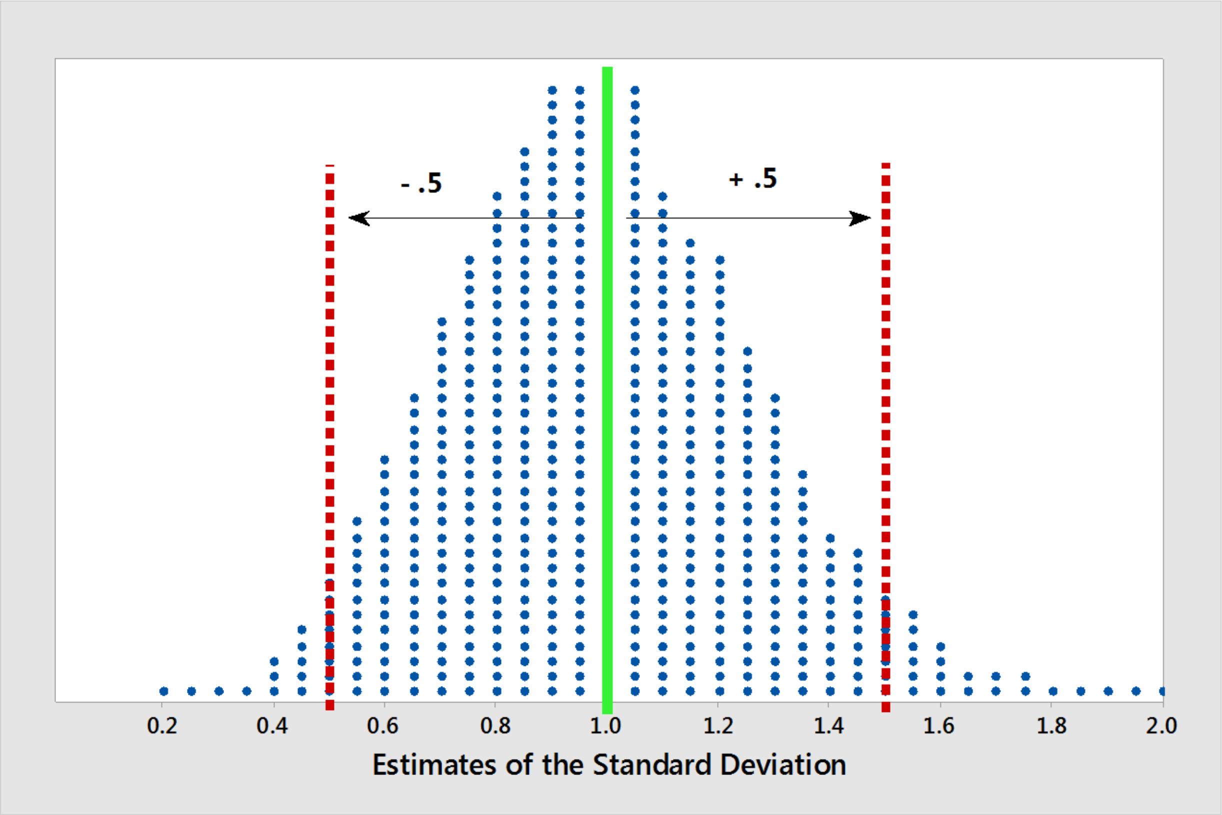
(5.31 , 6 , 6.69)

What about the standard deviation?

Confidence interval simulation when the standard deviation is not known for various sample sizes is shown by David M. Lane: http://onlinestatbook.com/2/estimation/ci_sim.html

JMC Data Experts

Distribution of 10,000 Estimates of the Standard Deviation



Conclusions

1. 95% of the estimates of the standard deviation between .5 and 1.5.
2. Your estimate from 8 samples will likely be within + / - 50% of the right answer.

How does this affect our perspective on process capability estimates and sample size? How does this affect our perspective on measurement system capability?

Source of Variation in Process Capability Estimates (Cpk)

This is not the problem

$$Cpk = \min \frac{[\ (\text{Upper Spec} - \text{Avg})\ \text{or}\ (\text{Avg} - \text{Lower Spec})\ ]}{(\ 3 * \text{Standard Deviation})}$$

This is the main source of variation

To estimate the mean, 5 – 10 samples might be acceptable. Because of the variability in the standard deviation, 30 samples or more are used to estimate Cpk

JMC Data Experts

Estimating Measurement System Capability

$$\sigma^2_{Total} = \sigma^2_P + \left(\left(\sigma^2_O + \sigma^2_{PO}\right) + \sigma^2_\varepsilon\right)$$

$$\sigma^2_{Total} = \sigma^2_P + \left(\sigma^2_{reproducibility} + \sigma^2_{repeatability}\right)$$

$$\sigma^2_{Total} = \sigma^2_P + \sigma^2_{measurement}$$

$$\boxed{\%Study\ Variation = \frac{\sigma_{measurement}}{\sigma_{total}} \times 100\%}$$

Rarely Can Trust This

$$\boxed{\%Tolerance = \frac{6 \times \sigma_{measurement}}{USL - LSL} \times 100\%}$$

My First Choice

$$\boxed{\%\ Process\ Variation = \frac{\sigma_{measurement}}{\sigma_{Historical}} \times 100\%}$$

My Second Choice

Expanded Gage R & R Overview

Overall Process Variation

Part-to-Part Variation

1

Measurement System Variation

Within Gage Variation

2

Gage to Gage Variation

3

Part * Gage Interaction

4

Variation due to Measurement Procedure

(Repeatability)

Operator * Gage Interaction

7

Oper-to-Oper Variation

5

Operator * Part Interaction

6

(Reproducibility)

Sampling Plan for an Expanded Gage R & R Study

Standard Gage Study:

10 Parts X 3 Operators X 2 Repeats

Measurement System Analysis, 4rd Edition, 2010, AIAG

Expanded

X 3 Gages

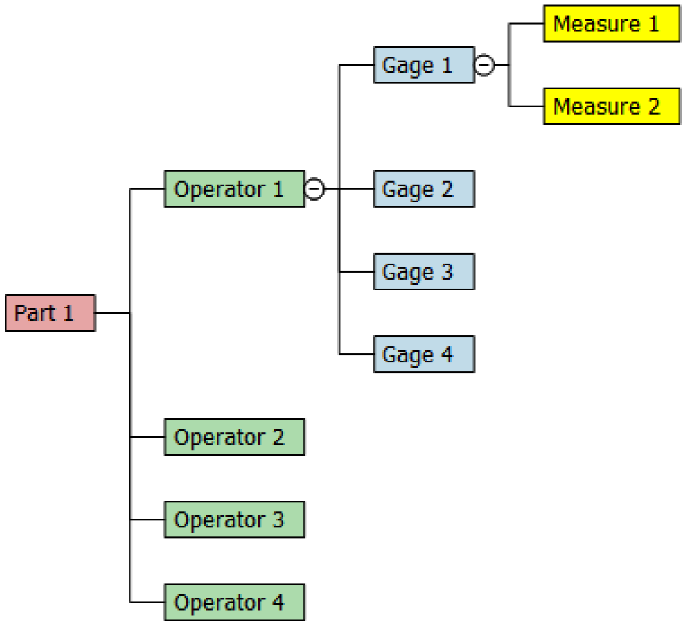
= 180 Measurements*

$$\sigma^2_{Total} = \sigma^2_P + \left(\left(\sigma^2_O + \sigma^2_{OG} + \sigma^2_{PO}\right) + \left(\sigma^2_G + \sigma^2_{PG}\right) + \sigma^2_\varepsilon\right)$$

$$\sigma^2_{Total} = \sigma^2_P + \sigma^2_{measurement}$$

$$\boxed{\%\,\text{Process Variation} = \frac{\sigma_{measurement}}{\sigma_{historical}} \times 100\%}$$

$$\boxed{\%\,Tolerance = \frac{6 \times \sigma_{measurement}}{USL - LSL} \times 100\%}$$



Sampling Plans for Expanded Gage R & R Studies, Lou Johnson and Daniel Griffith https://www.youtube.com/watch?v=KtGhnimE6Qw

How the Simulation Works

In a gage R&R study, we have the following model:

	Mean	Part	Operator	Gage	Repeatability
$Y_{ijkr} =$	$\mu +$	$P_i +$	$O_j +$	$G_k +$	ε_{ijkr}

1) Specify a SD for Part, Operator, Gage and Error -----------
2) Specify # of Samples for Part, Operator and Gage

Reset SD's and Sample Size

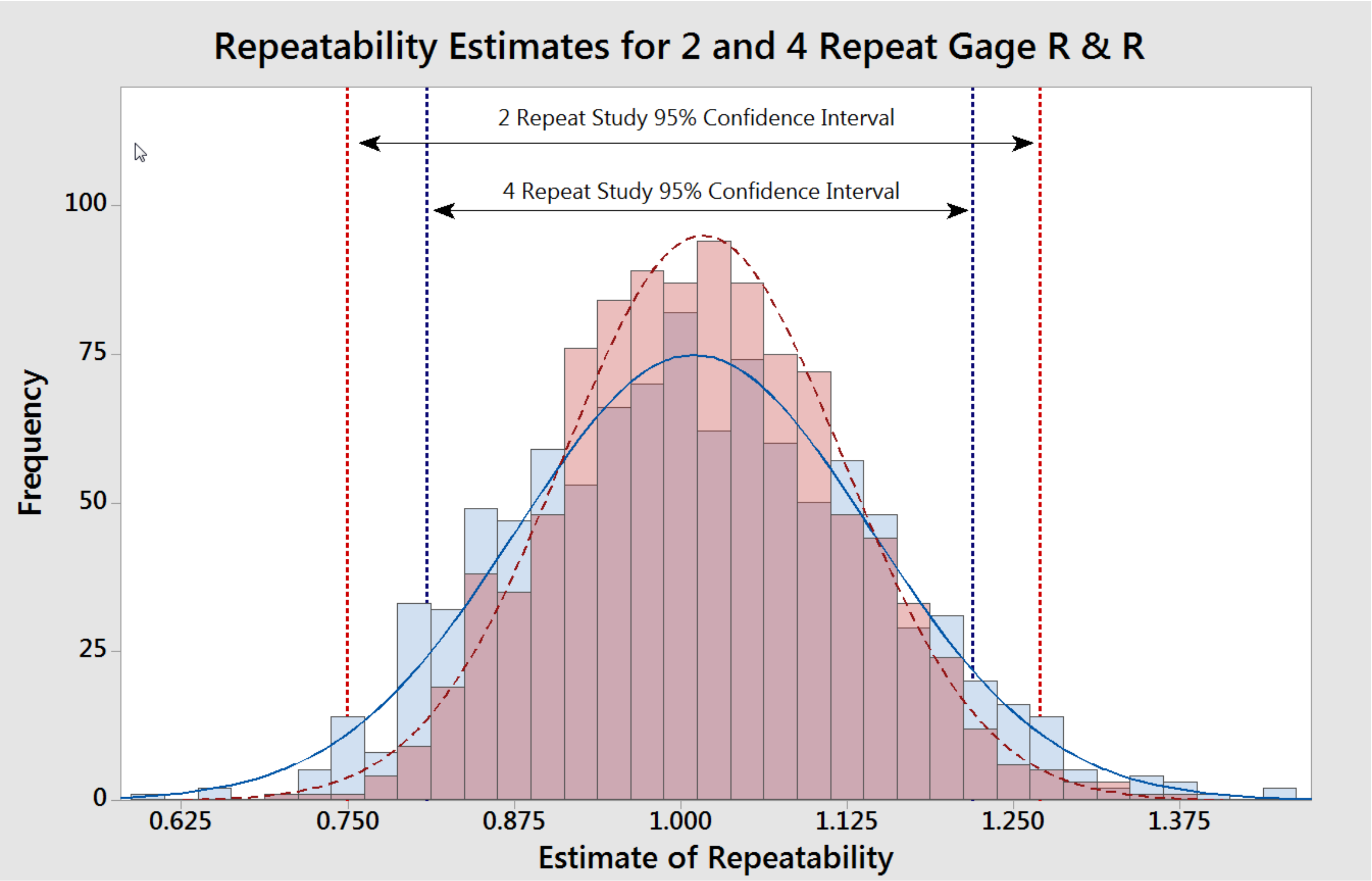
Simulate Data

C5	C6-T	C7-T	C8
Part	Gage	Operator	Response
1	Gage D	Mike	10.3
3	Gage D	Ann	9.9
4	Gage A	Mike	10.1
3	Gage B	Lou	9.8
1	Gage C	Ann	10.1
4	Gage B	Mike	10.2

Analyze Gage R&R –
Calculate Variance Components

σ_O^2
σ_P^2
σ_G^2
σ_e^2

95% Confidence Interval for Repeatability – 4 vs 2 Repeats



Compare

10 Parts X 3 Oper X 2 Repeats = 60 readings

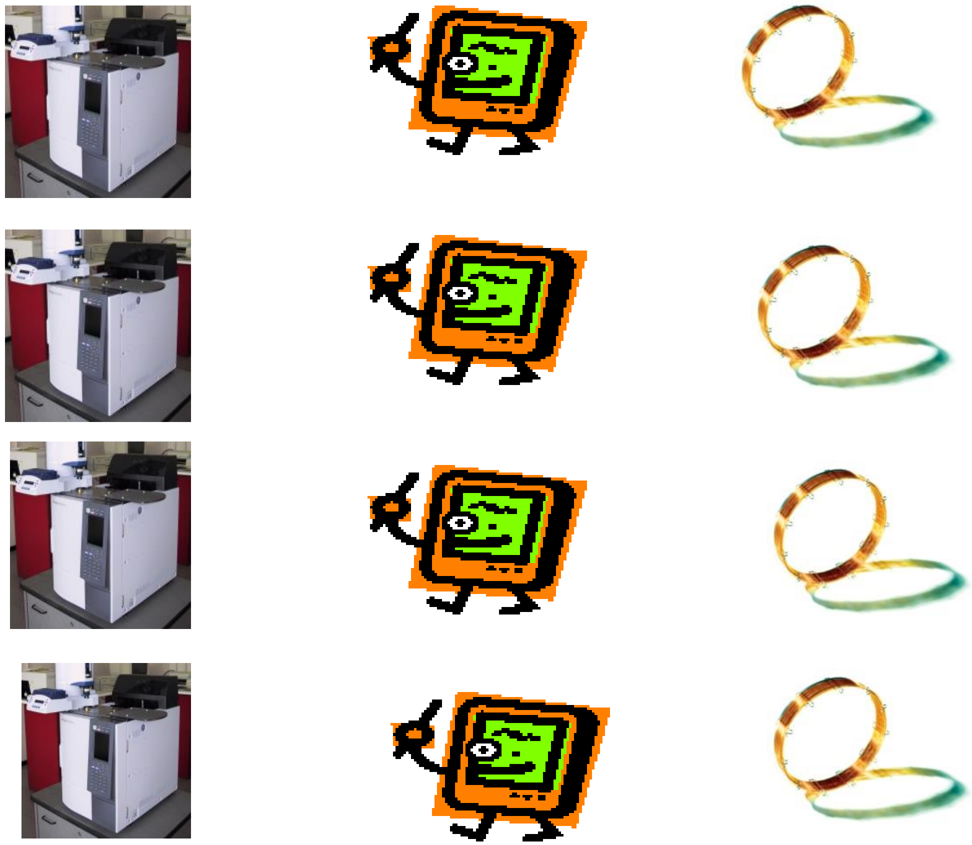
5 Parts X 3 Oper X 4 Repeats = 60 readings

95% CI on Repeatability (2 Repeats)
(1.27 - .75) = .52

95% CI on Repeatability (4 Repeats)
(1.22 - .81) = .41

21% lower variability in your estimate of repeatability just by changing the sampling plan.

Sampling Plan for Chromatograph Peak Height Expanded Study



4 Chromatographs (Gage)

4 Techs (Operator)

4 Columns (Part)

- **4 Columns used to run the 4 Chromatographs and 4 Techs in parallel**
- **2 Repeats**
- **96 measurements**

This turns out to be the best overall sampling plan for an expanded study when the variation for each component can not be estimated.

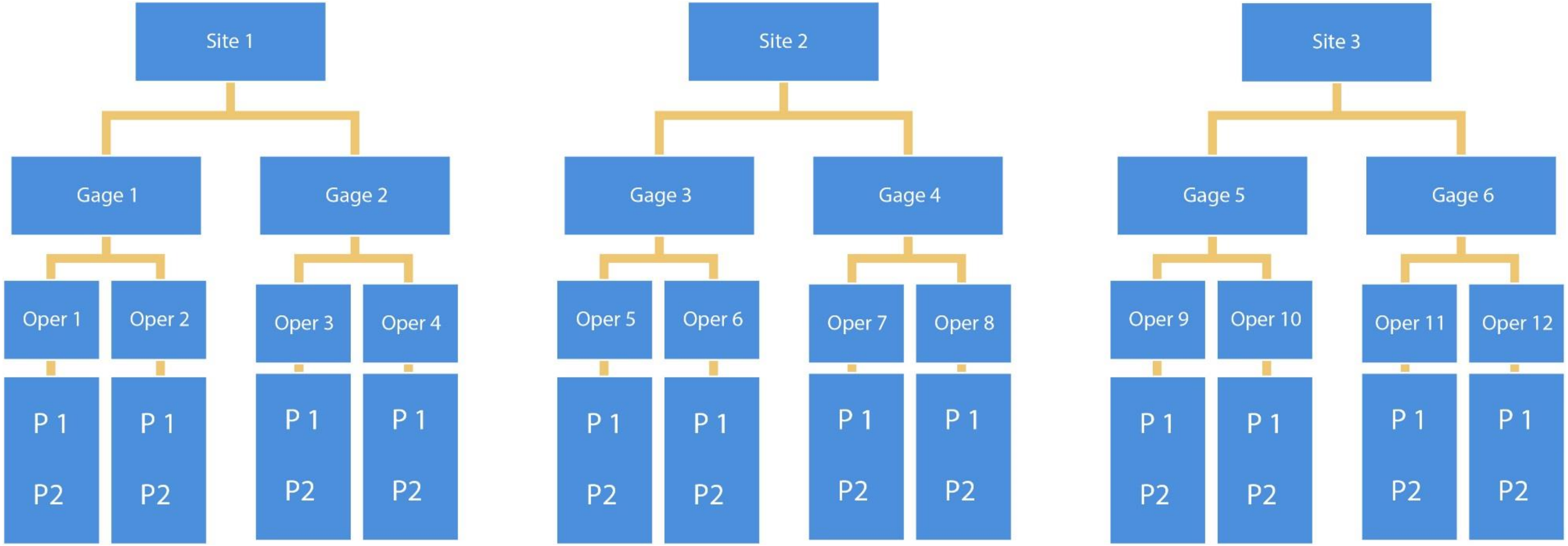
JMC Data Experts

Sample Size for Analysis of Variance Study

A syringe needle manufacturer measures force – till – failure on an Instron mechanical properties gauge.

Product is made and measured at three manufacturing sites. The third manufacturing site does not always meet specifications whereas the other two do. The goal is to determine which sources of variation (process /site, gage, operator, within lot) are the strongest contributors to the variation and therefore out-of-specification measurements.

Sampling Plan to Determine Sources of Variation with Nesting



10 repeat measurements (needles) of 2 Parts crossed with 12 Operators nested in 6 Gages nested in 3 Locations.

Nesting makes it hard to separate sources of variation. More samples (needles) helps, but how many more?

JMC Data Experts

Results of the Study -- Focus on Location and Repeatability

Variance Components

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.0979	12.82
Repeatability	0.8338	10.91
Reproducibility	0.1459	1.91
Location	0.0134	1.76
Operator(Location)	0.0009	0.12
Instron(Location)	0.0002	0.03
Part-To-Part	0.6661	87.18
Part	0.6661	87.18
Total Variation	0.7641	100.00

Repeatability `is the largest source of variation, but this contains the needle – to – needle variation as well as the measurement variation.

Location is the next largest source of variation, eclipsing Operator and Instron.

So this study was successful, but did we need to use all those samples? 10 per condition?

JMC Data Experts

Sample and Resample the Dataset with Different Sample Sizes

Analysis of Variance Model:

$$\underline{\text{Mean}}\quad \underline{\text{Loc}}\quad \underline{\text{Part}}\quad \underline{\text{Operator}}\quad \underline{\text{Gage}}\quad \underline{\text{Repeatability}}$$

$$Y_{sijkr} = \mu + L_s \quad P_i + \quad O_j + \quad G_k + \quad \varepsilon_{ijkr}$$

1) Specify the size of a random sample - 2,4,6, or 8 data pts.
2) Randomly select that many points from the 10 in the data

Select a New Sample

Sample the Data Accordingly

Repeat 5000 times
Histogram of Answers

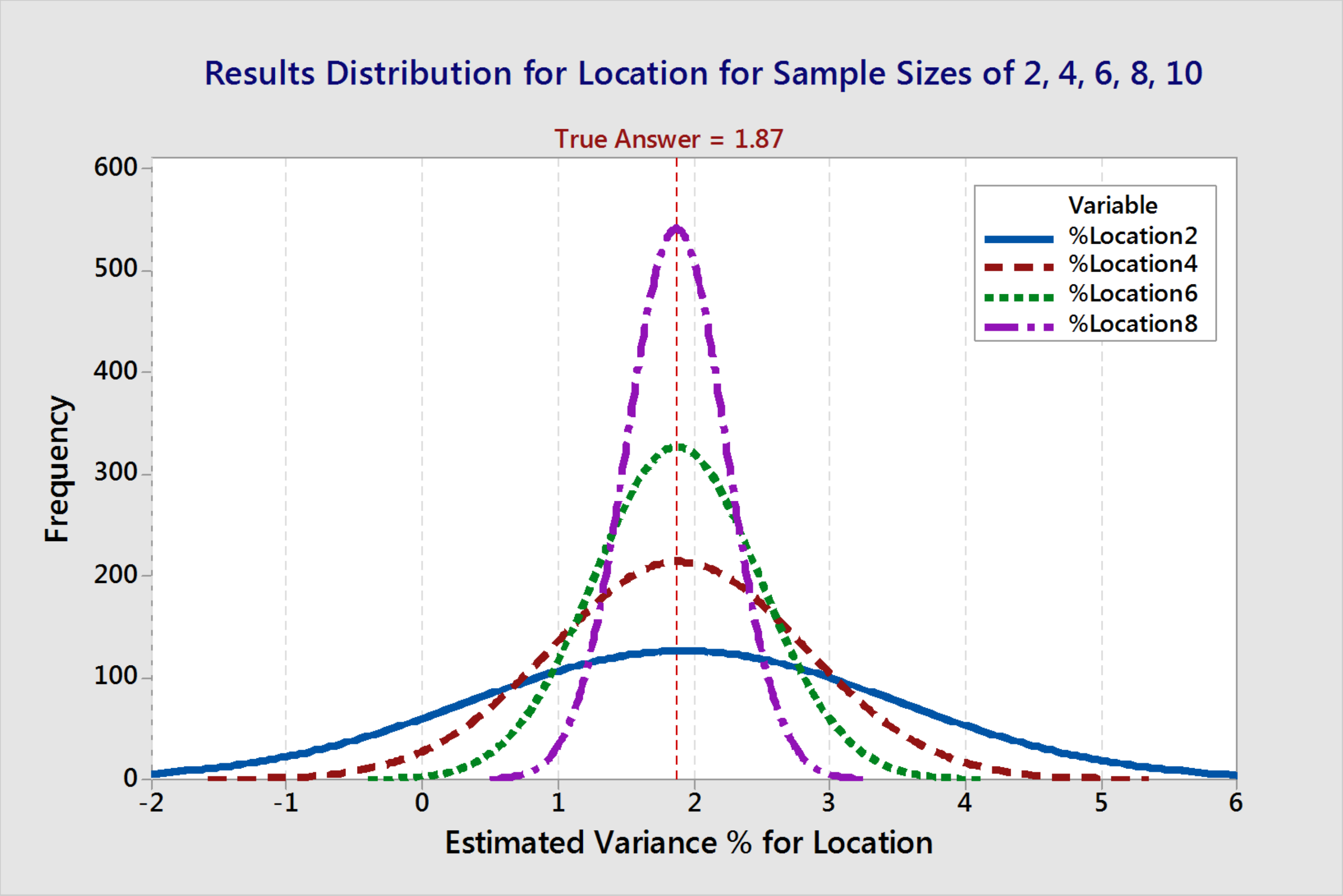
C5 Part	C6-T Gage	C7-T Operator	C8 Response
1	Gage D	Mike	10.3
3	Gage D	Ann	9.9
4	Gage A	Mike	10.1
3	Gage B	Lou	9.8
1	Gage C	Ann	10.1
4	Gage B	Mike	10.2

Analyze Dataset –
Calculate Variance Components

$$\boxed{\begin{array}{c}\sigma_O^2\\ \sigma_L^2\\ \sigma_G^2\\ \sigma_e^2\end{array}}$$

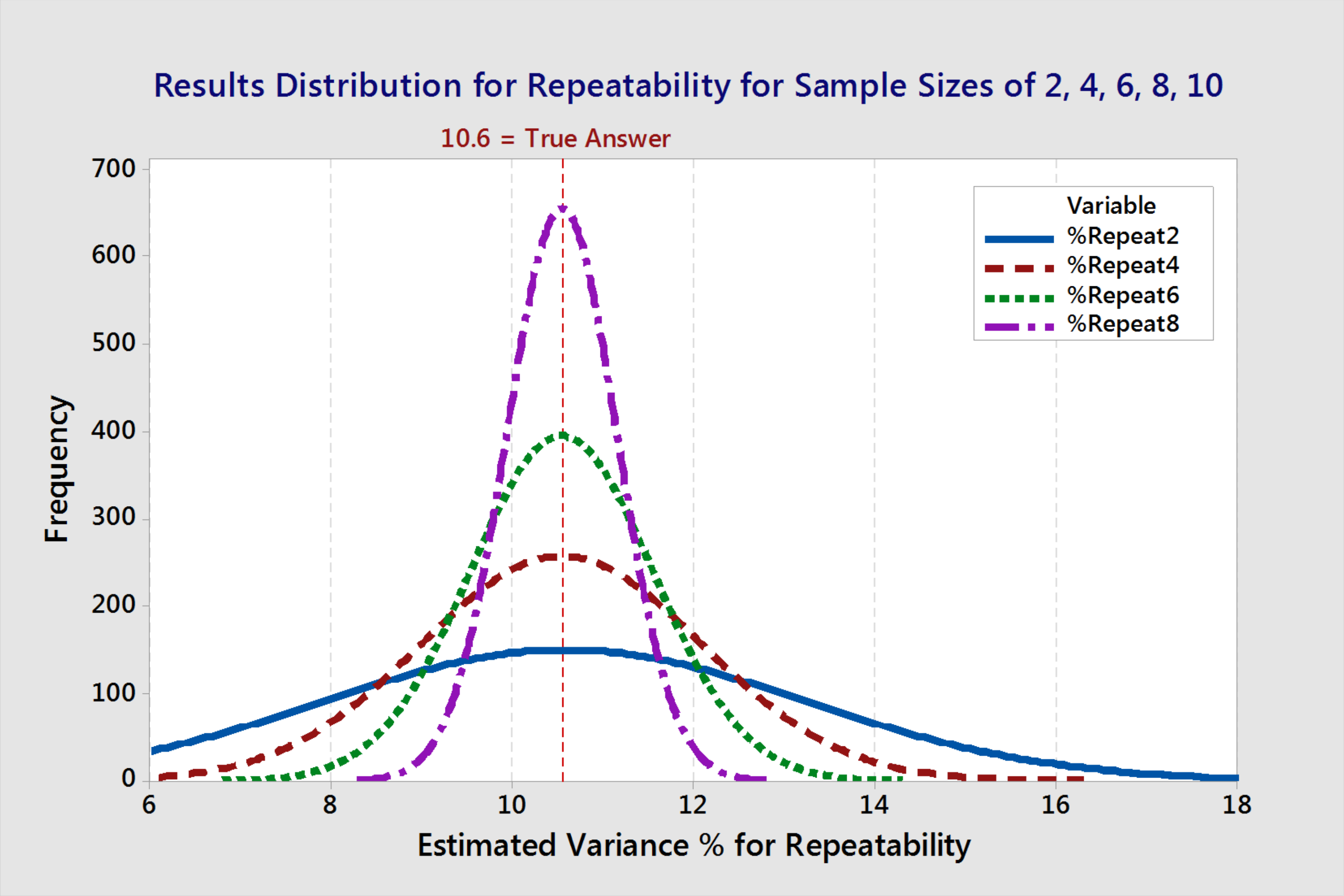
Simulations by Cathy Akritas at Minitab, Inc cakritas@minitab.com

Results of the Simulations

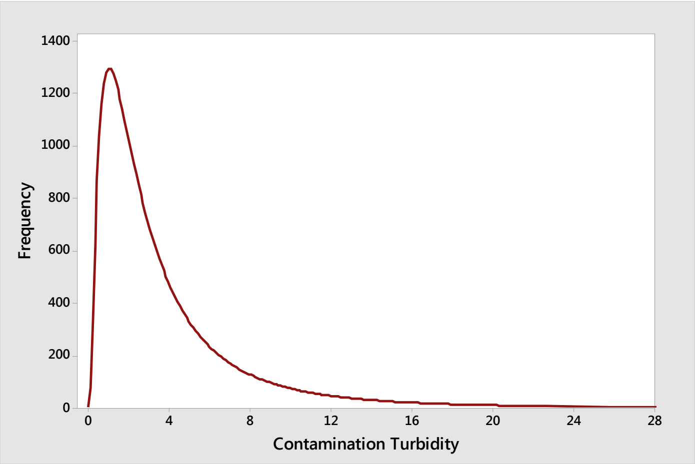


JMC Data Experts

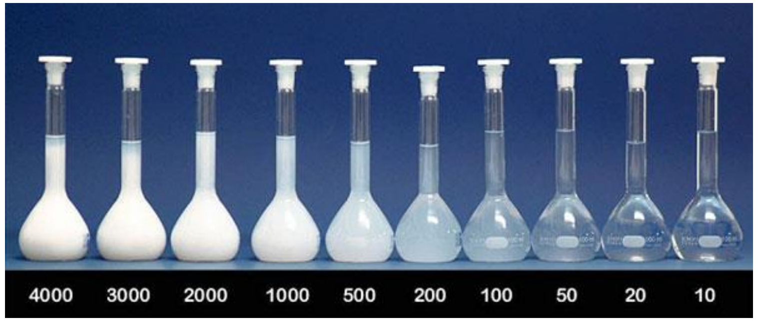
Results of the Simulations



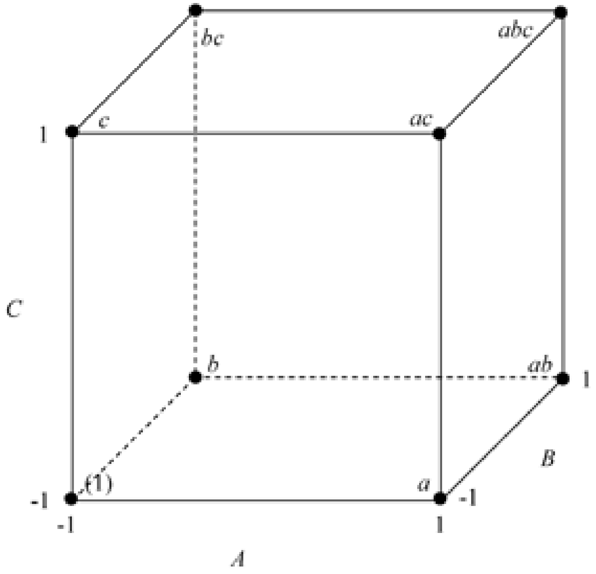
Sample Size - 3 Factor Factorial with a Non-Normal Covariate



+



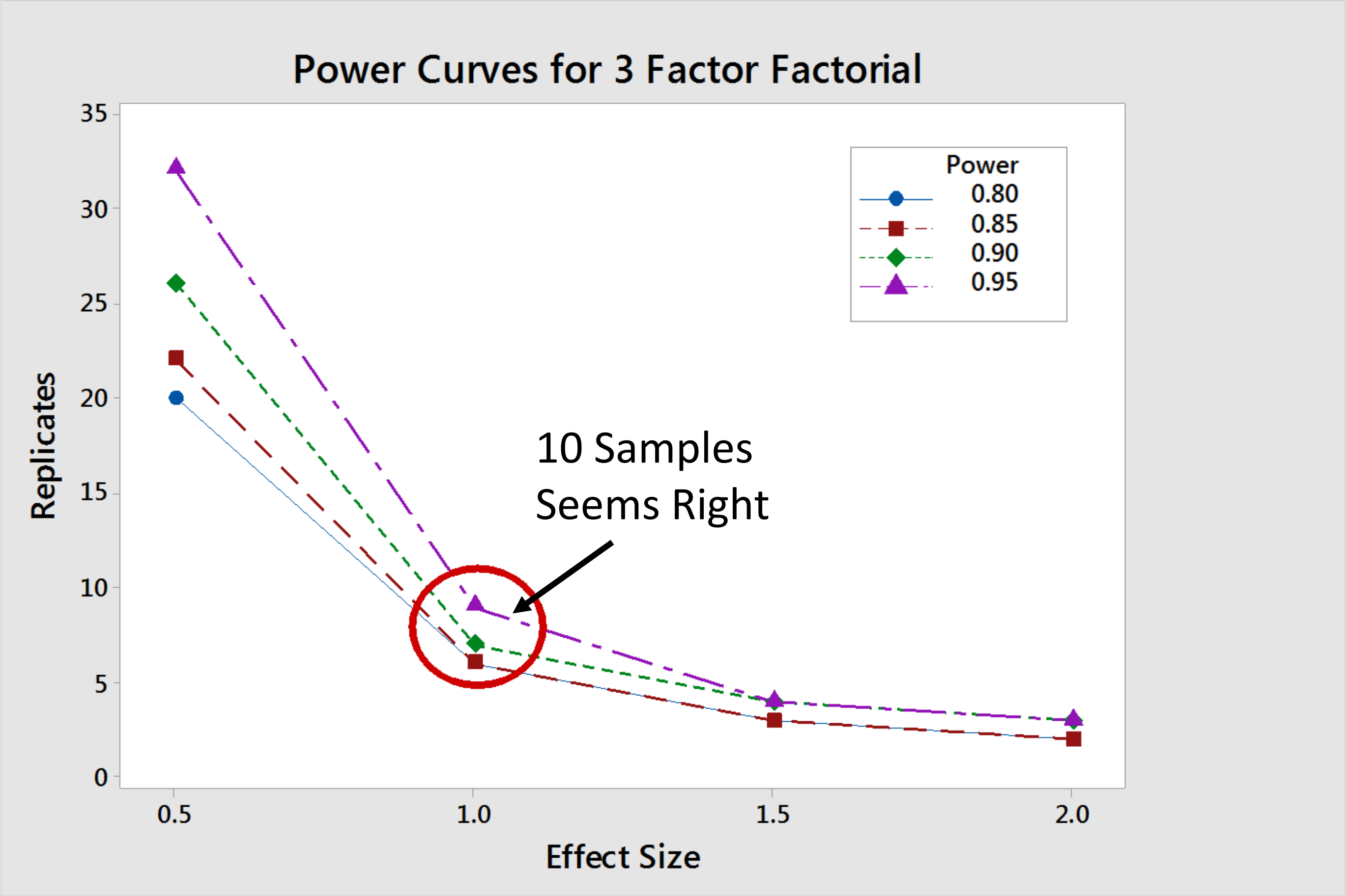
Incoming Contaminant
Random Lognormal



Three Factor Cleaning
Process Experiment

Clean Product

Sample Size for a 3 Level Factorial with a Non-Normal Covariate



The power calculation can seem reasonably straight forward. 10 samples per treatment is reasonable.

But the response:

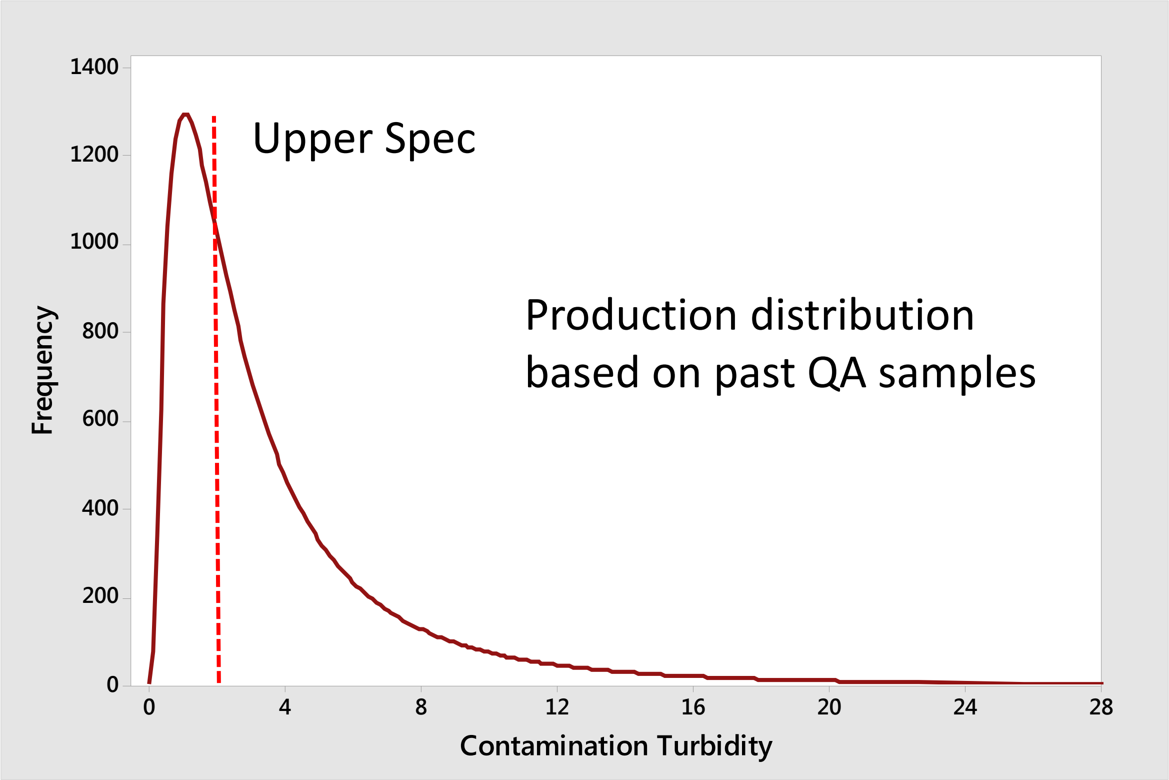
Turbidity Before - Turbidity After

is function of **Turbidity Before**.

1.1 - .7 = .4 and
16.4 – 16.0 = .4
have completely different meanings even though they are equal.

Even with random sampling , enough samples are needed to present each condition with the same distribution of incoming contamination.

Sample Size for a 3 Level Factorial with a Non-Normal Covariate

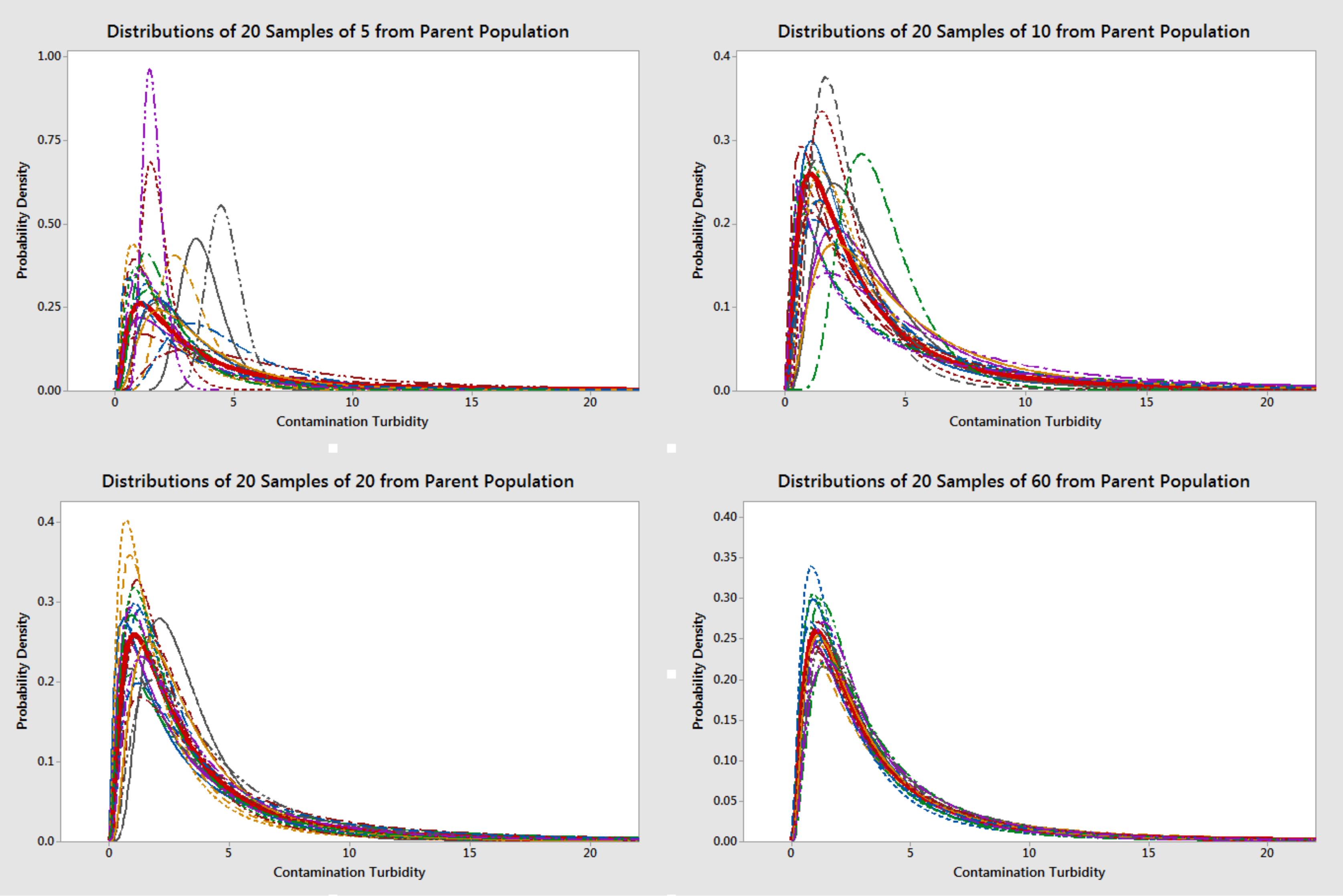


How many samples must be taken from this distribution before the samples will duplicate the parent distribution with reasonable consistency.

Randomly sample 5, 10, 15, 20, …… etc. , twenty times and plot their distribution with the parent.

JMC Data Experts

Sample Size to Consistently Replicate the Distribution



JMC Data Experts

Application of Simulation to Sample Size Calculations and Teaching Statistical Concepts

Questions?

Presented by:
Lou Johnson
lou@jmcdataexperts.com

JMC Data Experts

www.jmcdataexperts.com