DATA-DRIVEN INSIGHTS WITH MONTE CARLO SIMULATION

Charles Cox

March 1 – 4, 2021

Learning Objectives

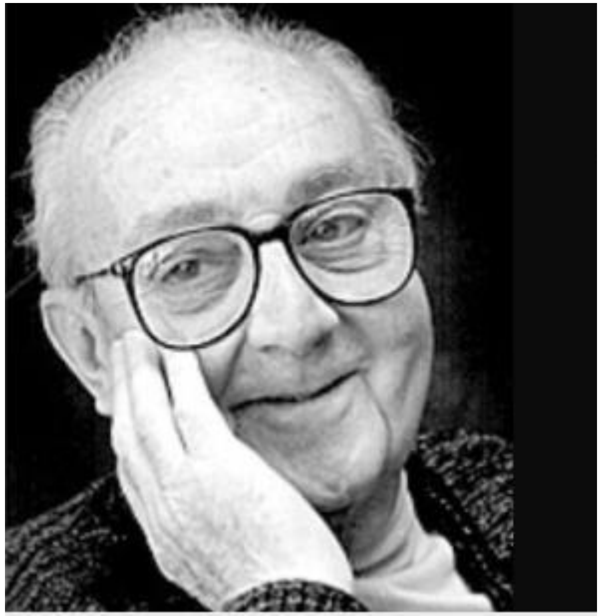
In this session you will:

- Learn the benefits of Monte Carlo simulation and how to take your process understanding to the next level
- Learn how to create an initial approximation of any process step’s duration or cycle time
- Learn how to run the initial Monte Carlo **simulation**, get initial results and determine what to do next
- Learn how to do **Parameter Optimization** on input distributions’ parameter values
- Learn to do **Sensitivity Analysis** to determine where to target input improvements that will give the greatest overall improvement in the process’ results

Intro to Simulation and Models

- Uncertainty, ambiguity, and variability: these go hand in hand with the inherent risk that is a part of all business decisions
- Is there a way to understand and quantify this risk so that organizations can make better business decisions?
- How can we get to the next level of insight? Create a model

Intro to Models and Modeling



George E. P. Box

"…all models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind…"

Empirical Model-Building and Response Surfaces, 1987

Deterministic vs Random System

- A deterministic system is one in which no randomness is involved in the future states of the system. It always produces the same output from the same given starting conditions
- When randomization is to be accounted for, as in a stochastic system, it’s useful to model that system
- Putting more information ‘in’, gets more information ‘out’:
  - Characterize the inputs’ randomness with:
    - i) distributions and,
    - ii) the distributions’ parameters, and the
    - Transfer function, $F(x) = Y$
- Determine if the range of outcomes is acceptable or not…and if not, determine next steps

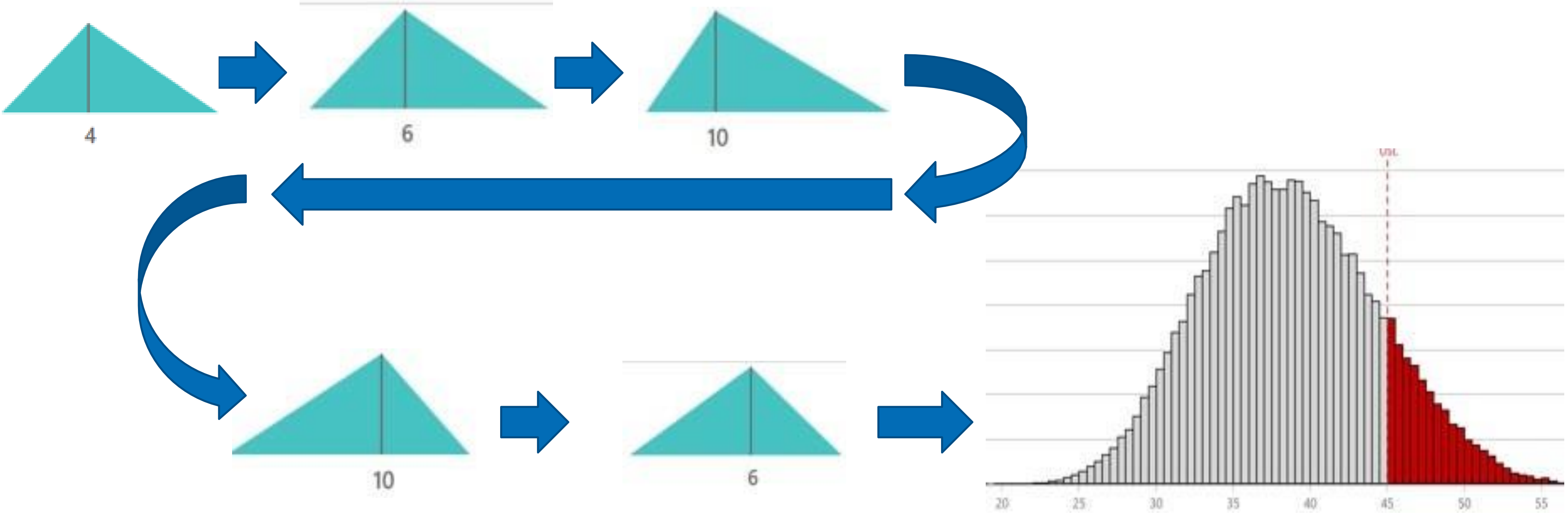
Which approach gives more information?

- Is the point value output acceptable?   USL = 45 business days

#1 → #2 → #3 → #4 → #5 → Output

4 + 6 + 10 + 10 + 6 = 36

- Is range of outcomes acceptable?   ~14% > 45 business days



Applying Monte Carlo Simulation

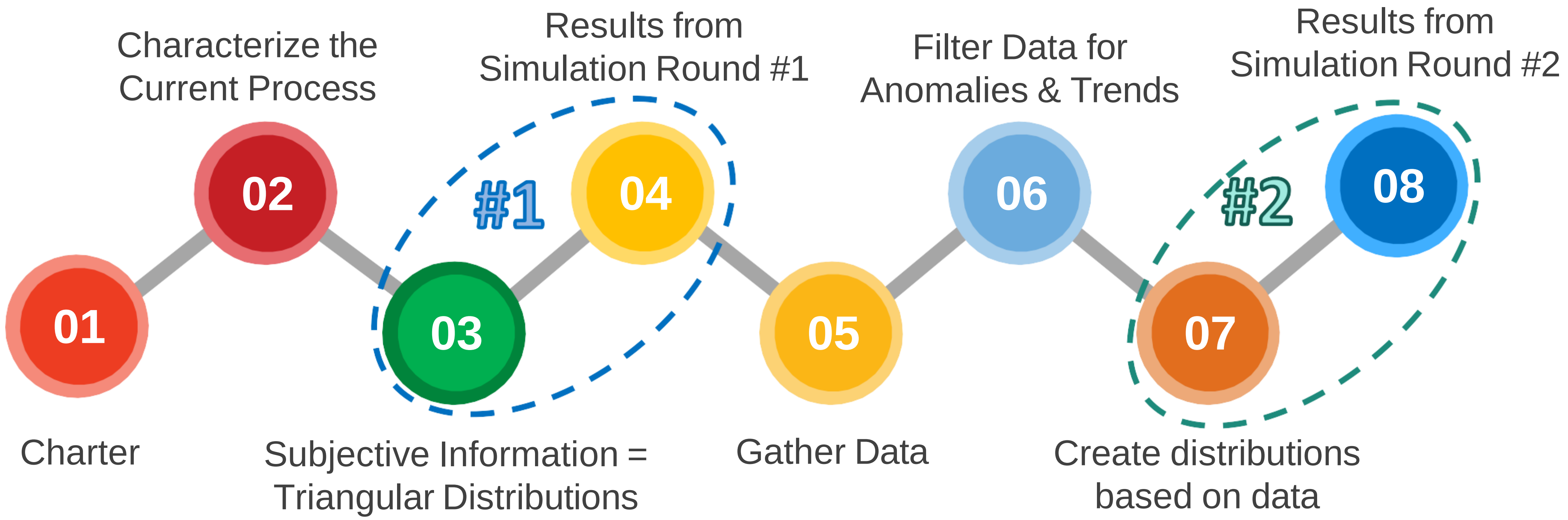
- In the design world, in transactional and manufacturing operations, there are multiple ways of understanding and handling randomness. One of the most efficient ways is to use Monte Carlo simulation to model how the various distributions work together to understand what the impact on the final result will be.

Monte Carlo - History

- “**Monte Carlo simulations** are named after the gambling hot spot in Monaco, since chance and random outcomes are central to the modeling technique, much as they are to games like roulette, dice, and slot machines.”

– Investopedia.com

Applying Monte Carlo to a Specific Case



- Charter – Goal, In-scope, Out-of-Scope, Beginning and Ending points of the process, Data Sources, Resources
- Characterize – Research Minimum, Maximum and Most Likely Cycle Time for each step in the Current Process
- In Round #1, use Triangular Distributions based on subjective information for doing the Monte Carlo simulation
- Begin gathering variable data for Round #2, and filter data
- In Round #2, to do the Monte Carlo Sim, use the Distributions indicated by the data that's been collected after Round #1

Cycle Time Data

- Working in the real world of process improvement, a team will need to establish distributions for the cycle times of each of the tasks in the process' Scope of Work. Often at the start of a new improvement project for a transactional process, there will be little data available, so some information gathering will be necessary

Cycle Time Data (cont.)

- Interview people working in each of the tasks in the process and create triangular distributions based on their answers to questions such as:
  a) What is the longest this task has taken?
  b) What is the shortest time in which this task has been accomplished?
  c) What is the usual time to get this task finished?
- These 3 figures can serve as the basis for creating a triangular distribution

New Industrial Utility Account

Signing a new Client Inside Existing Grid

1. Initial application with funds: Open Account
2. Study on KWH **demand** for first 36 months
3. Study location on grid and **supply** point(s)
4. Report, final fees, rate structure
5. Acceptance and kick-off of Work Orders

The metric: All cycle times in business days (typical 5 business days per week, less any holidays)

Transactional Example – Initial Sim

- There are 5 steps in series; 5 inputs that give 1 output
- Note that the figures obtained are not from carefully timed tasks, but are the subjective opinions of the people actually working in the process on specific activities
- Assume no buffers between activities, so upon completion of first activity, the transaction moves immediately to the next step until all activities have been processed, and transaction is complete

TABLE 1 - Time in Business Days			
Step #	Shortest Time	Most Likely Time	Longest Time
1	2	4	7
2	2	6	12
3	4	10	25
4	3	10	14
5	2	6	9

Initial Simulation Set-up

Define Model

X Name	Distribution	Parameters: Lower	Mode	Upper	Preview
Step 1	Triangular	2	4	7	 4
Step 2	Triangular	2	6	12	 6
Step 3	Triangular	4	10	25	 10
Step 4	Triangular	3	10	14	 10
Step 5	Triangular	2	6	9	 6

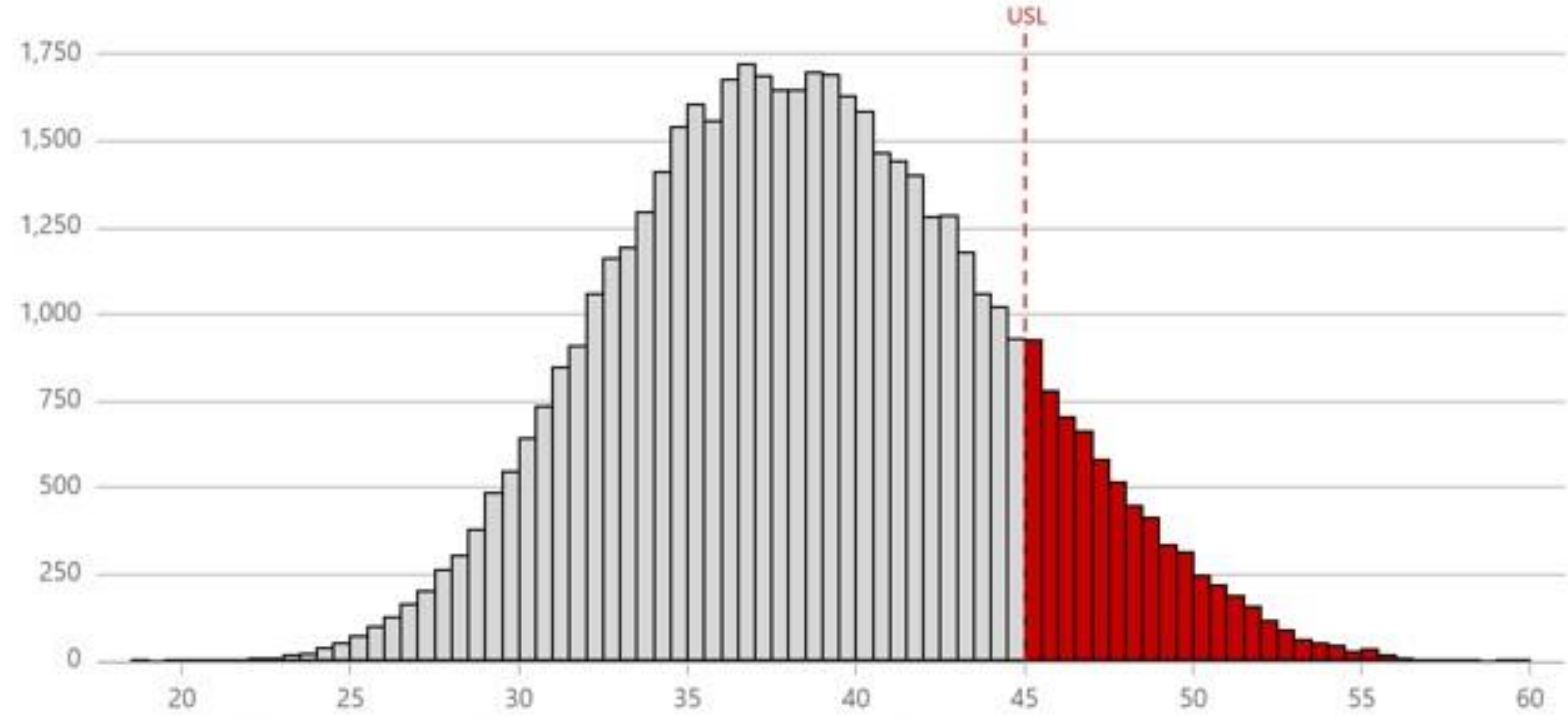
Initial Sim Equation and Model

Results of Initial Simulation

Simulation Results

Total Cycle Time #1

USL = 45



The simulation indicates that you can expect 14.08% of the *Total Cycle Time #1* values to fall outside of the specification limits. This corresponds to a Cpk of 0.3894. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.3894

% Out of Spec

14.08%

Summary Statistics

N	50,000
Mean	38.6481
Standard Deviation	5.6689

Transactional Example – Parameter Opt.

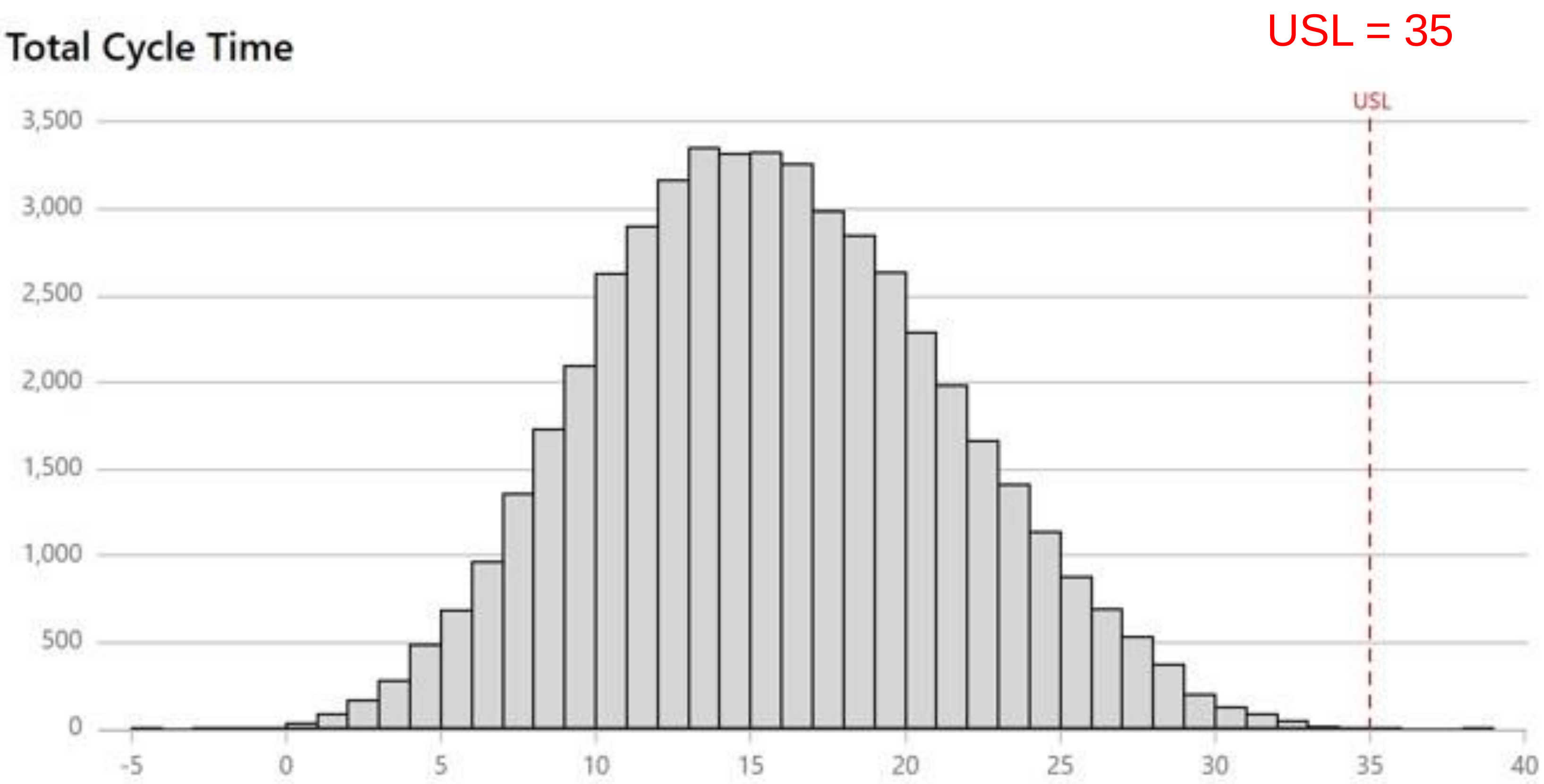
- Same 5 steps in series but with longest time reduced by 1 business day
- Will reducing the Longest Time by 1 business day, meet the goal of 7 business weeks (35 business days)?
- Assume everything else remains the same

TABLE 1 - Time in Business Days			
Step #	Shortest Time	Most Likely Time	Longest Time
1	2	4	6
2	2	6	11
3	4	10	24
4	3	10	13
5	2	6	8

Results After Parameter Optimization

Simulation Results

Total Cycle Time



The simulation indicates that you can expect 0.01% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 1.16. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

1.16

% Out of Spec

0.01%

Summary Statistics

N	50,000
Mean	15.7375
Standard Deviation	5.7022

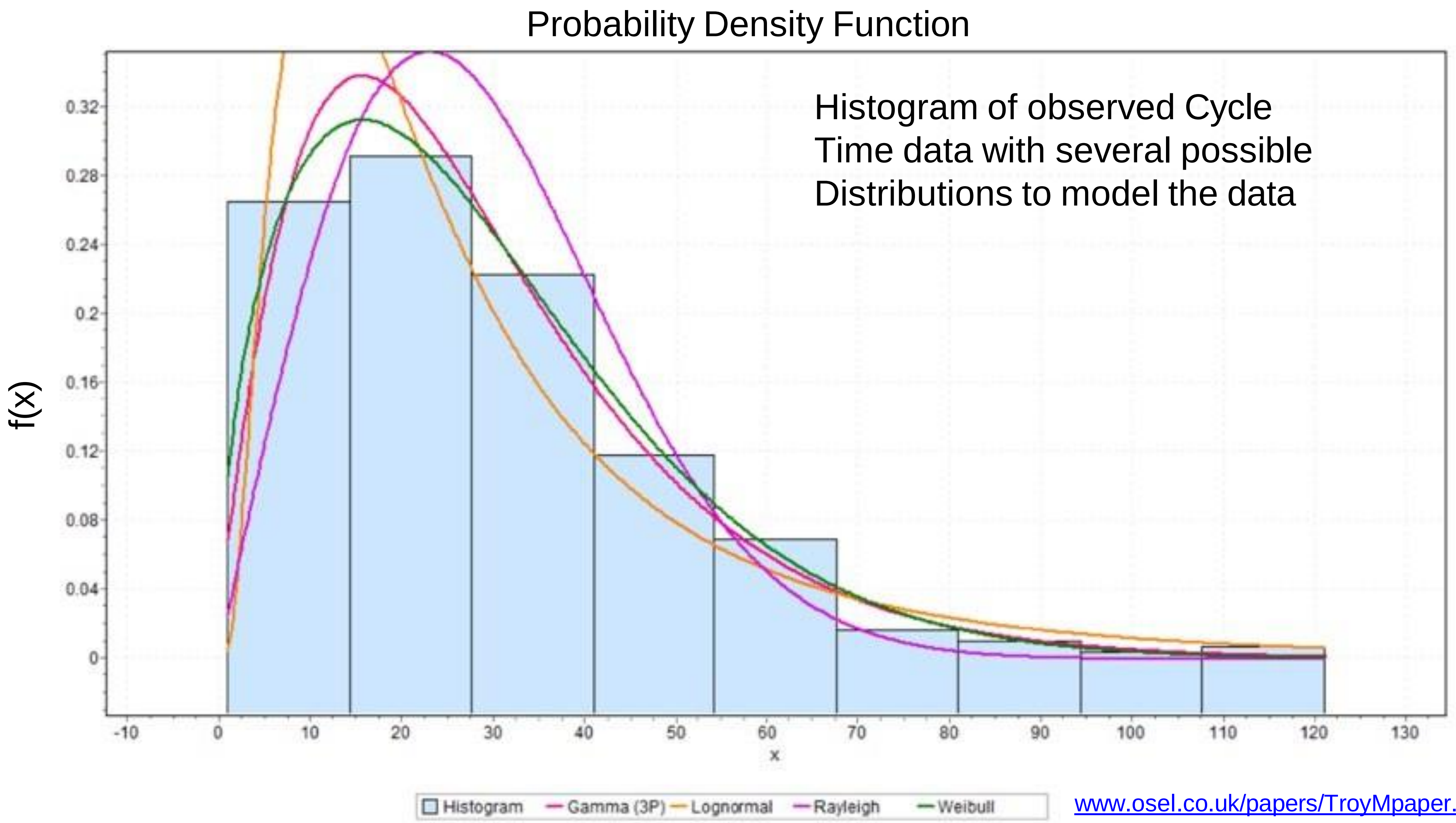
Goals: Current & Simulation of Improved (Parameter Optimization)

- Initial Simulation: Total Cycle Time is 9 business weeks with **USL = 45** Business Days
  Cpk = 0.39 and 14.08% Out of Spec
- Parameter Optimization (shorten Longest Time by 1 Business Day): Reduce 9 business weeks to 7 business weeks (~25% reduction)
  New USL = 35 Business Days
  Cpk = 1.16 and 0.01% Out of Spec

Gathering Actual Process Data

- In parallel with analyzing the Triangular Distributions based on subjective information, personnel began collecting actual cycle time data for their tasks
- In most instances, over 2 weeks to 2 months, it was possible to collect a statistically valid data set
- These data sets were the basis for continuous distributions such as the lognormal or Weibull distribution

Typical Transactional Data and Distributions

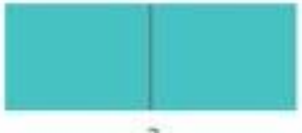
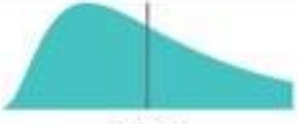


Histogram of observed Cycle Time data with several possible Distributions to model the data

www.osel.co.uk/papers/TroyMpaper.pdf

Simulation with Observed Data

Define Model

X Name	Distribution	Parameters			Preview
Step 1	Uniform	Lower: 2	Upper: 4		
Step 2	Weibull	Shape: 1.35	Scale: 4	Threshold: 2	
Step 3	Weibull	Shape: 1.7	Scale: 8.5	Threshold: 4	
Step 4	Weibull	Shape: 1.2	Scale: 4	Threshold: 3	
Step 5	Lognormal	Location: 0.4	Scale: 0.7		

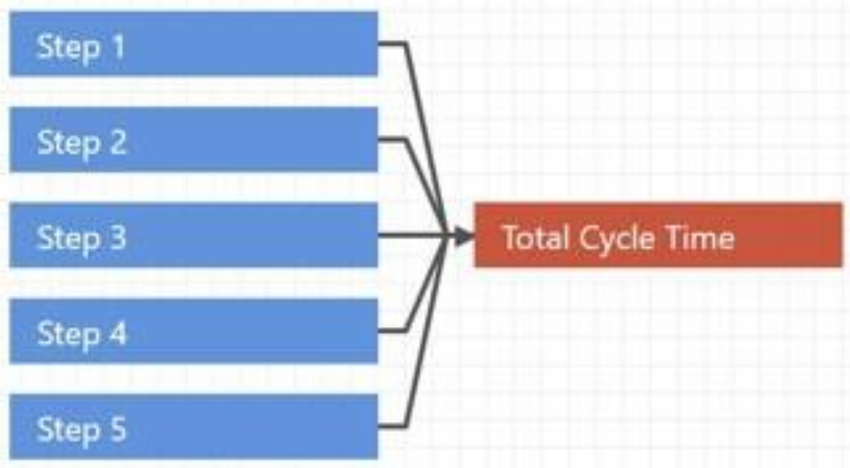
Observed Data Equation and Model

Y Name	Equation	Spec Limits (Optional) LSL	USL
Total Cycle Time	= Step 1+Step 2+Step 3+Step 4+Step 5		35

Add Another Y

Model

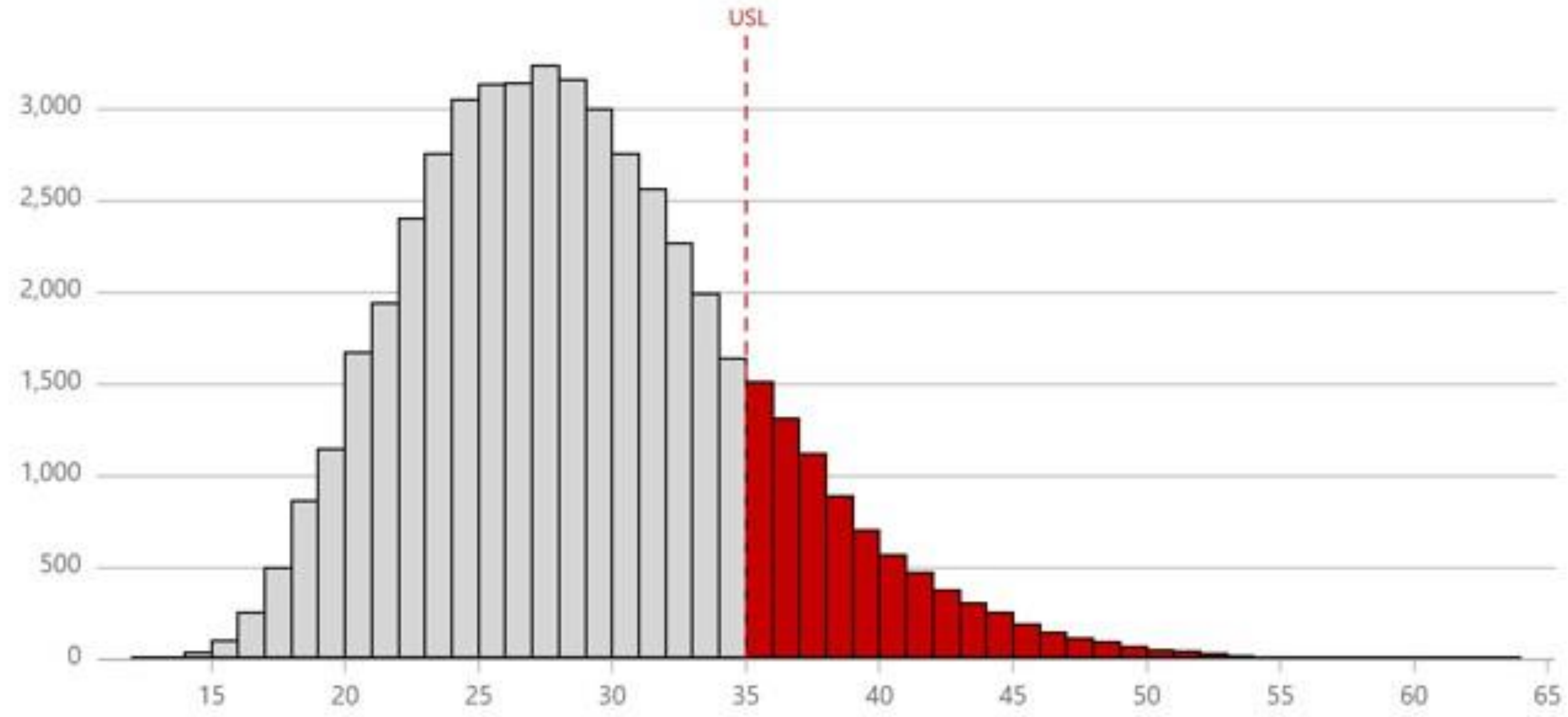
Before you run the simulation, use the diagram below to verify that the model is



Observed Data Sim Results

Simulation Results

Total Cycle Time



The simulation indicates that you can expect 16.69% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.274. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.274

% Out of Spec

16.69%

Summary Statistics

N	50,000
Mean	28.9079
Standard Deviation	6.4005

Observed Data Simulation Model

Model Assumptions

Inputs

Name	Distribution	Settings
Step 1	Uniform	(2; 4)
Step 2	Weibull	(1.35; 4; 2)
Step 3	Weibull	(1.7; 8.5; 4)
Step 4	Weibull	(1.2; 4; 3)
Step 5	Lognormal	(0.4; 0.7)

Outputs

Name	Equation
Total Cycle Time	Step 1+Step 2+Step 3+Step 4+Step 5

Observed Data Capability

Capability

Specification Limits

USL	35
LSL	—

DPMO

>USL	166,920
<LSL	—

Observed Performance

>USL	8,346
<LSL	—

Summary Statistics

N	50,000
Mean	28.9079
Standard Deviation	6.4005
Minimum	12.7725
Median	28.2384
Maximum	63.7271

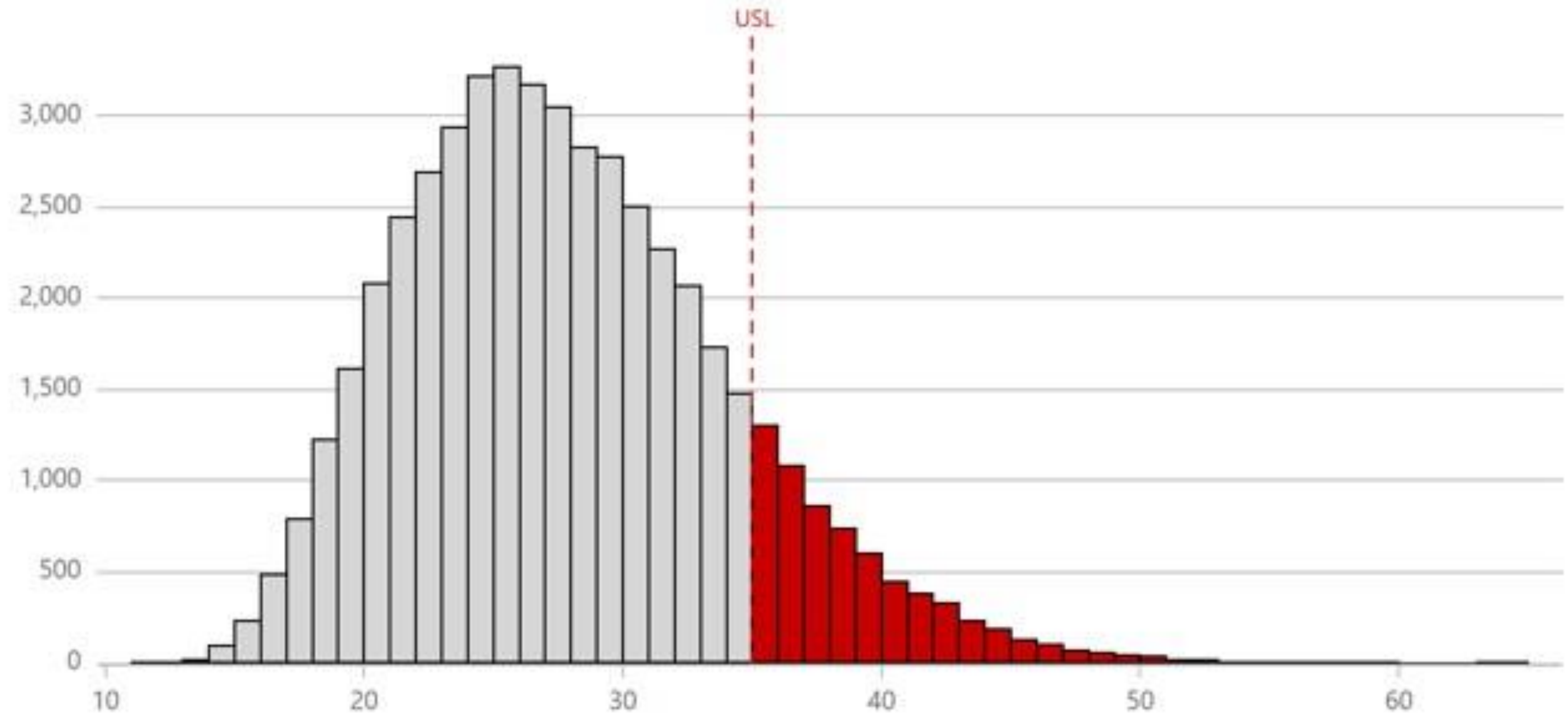
Percentiles

0.1^{st}	15.1319
0.5^{th}	16.4953
1^{st}	17.2299
5^{th}	19.6714
10^{th}	21.2271
90^{th}	37.4354
95^{th}	40.476
99^{th}	46.7449
99.5^{th}	49.0812
99.9^{th}	53.6959

Parameter Optimization Results

Parameter Optimization Results

Total Cycle Time



Using the new input settings, the simulation indicates that you can expect 13.66% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.311. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.311

% Out of Spec

13.66%

Summary Statistics

N	50,000
Mean	27.9077
Standard Deviation	6.4019

Parameter Optimization Results

Assumptions

Optimization Goal: Minimize the % Out of Spec of Total Cycle Time

Inputs

Name	New Settings	Search Range		Previous Settings	Distribution
Step 1	(1.0013; 3.0013)	Low: 2	High: 3	(2; 4)	Uniform
Step 2	(1.35; 4; 2)	Low: —	High: —	(1.35; 4; 2)	Weibull
Step 3	(1.7; 8.5; 4)	Low: —	High: —	(1.7; 8.5; 4)	Weibull
Step 4	(1.2; 4; 3)	Low: —	High: —	(1.2; 4; 3)	Weibull
Step 5	(0.4; 0.7)	Low: —	High: —	(0.4; 0.7)	Lognormal

Outputs

Name	Equation
Total Cycle Time	Step 1+Step 2+Step 3+Step 4+Step 5

Parameter Optimization Results

Capability

Specification Limits	
USL	35
LSL	—

DPMO	
>USL	136,620
<LSL	—

Observed Performance	
>USL	6,831
<LSL	—

Summary Statistics

N	50,000
Mean	27.9077
Standard Deviation	6.4019
Minimum	11.7585
Median	27.1939
Maximum	64.6741

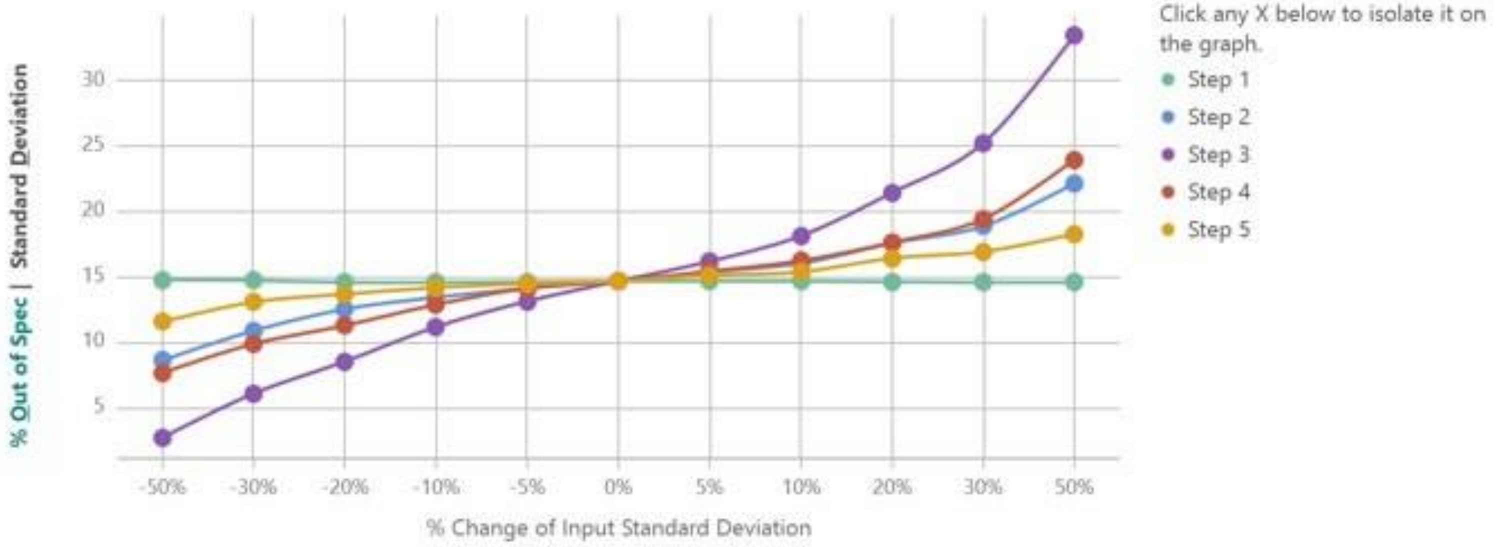
Percentiles

0.1st	14.1694
0.5th	15.5678
1st	16.3268
5th	18.7229
10th	20.2371
90th	36.4189
95th	39.5026
99th	45.6959
99.5th	48.2246
99.9th	53.0135

Sensitivity Analysis Choices

Sensitivity Analysis for Total Cycle Time

Use the graph to understand how changing the standard deviation of an input (X) affects the output (Y). Look for inputs that may improve the capability (sloped lines) and inputs that may allow less stringent tolerances (flat lines).



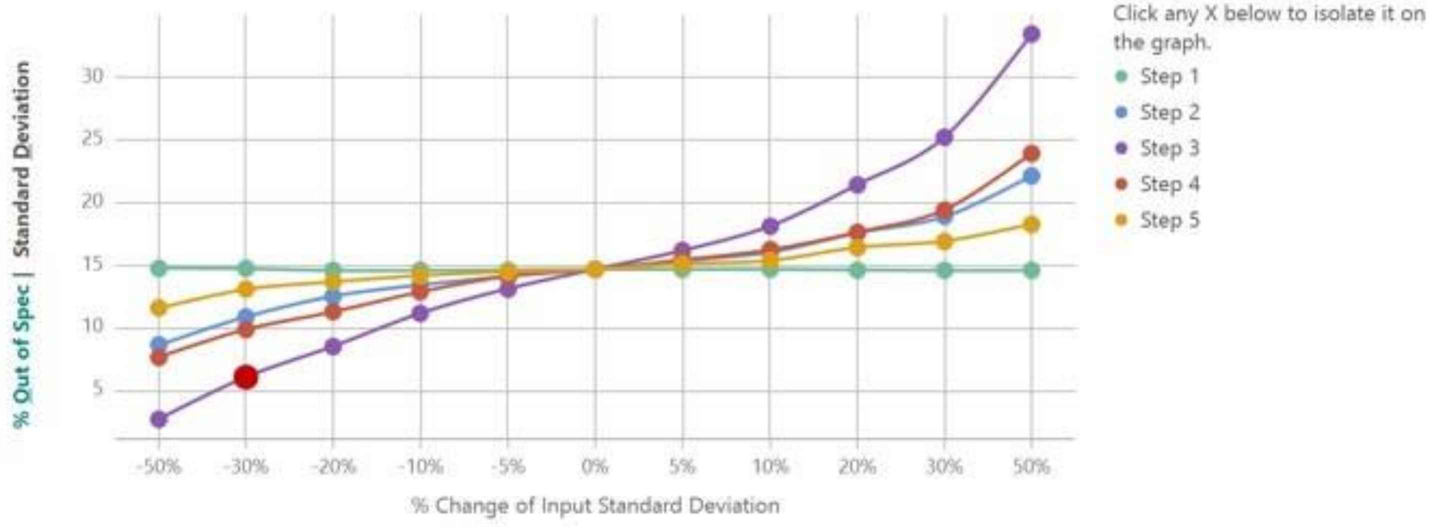
To change an input setting, click a point on the graph or choose a % change value from the table below.

Inputs (X)	Current St Dev	% Change	Proposed St Dev
Step 1	0.57735	0%	—
Step 2	2.7463	0%	—
Step 3	4.592	0%	—
Step 4	3.1489	0%	—
Step 5	1.5156	0%	—

Sensitivity Analysis 1st Choice

Sensitivity Analysis for Total Cycle Time

Use the graph to understand how changing the standard deviation of an input (X) affects the output (Y). Look for inputs that may improve the capability (sloped lines) and inputs that may allow less stringent tolerances (flat lines).



To change an input setting, click a point on the graph or choose a % change value from the table below.

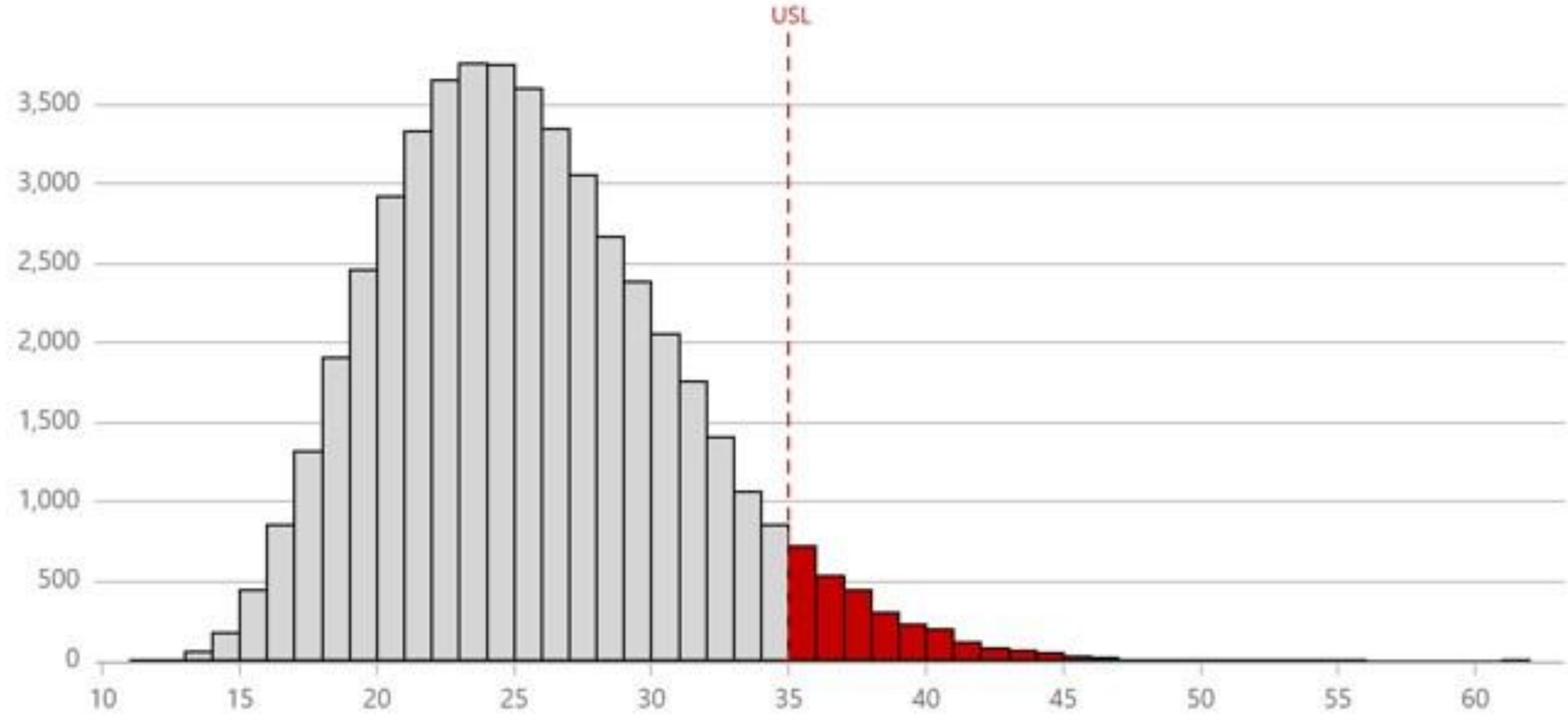
Inputs (X)	Current St Dev	% Change	Proposed St Dev
Step 1	0.57735	0%	—
Step 2	2.7463	0%	—
Step 3	4.592	-30%	3.2144
Step 4	3.1489	0%	—
Step 5	1.5156	0%	—

Reduce Step 3 by 30%

Sensitivity Analysis 1st Results

Sensitivity Analysis Results

Total Cycle Time



Using the new standard deviation values, the simulation indicates that you can expect 5.97% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.4438. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.4438

% Out of Spec

5.97%

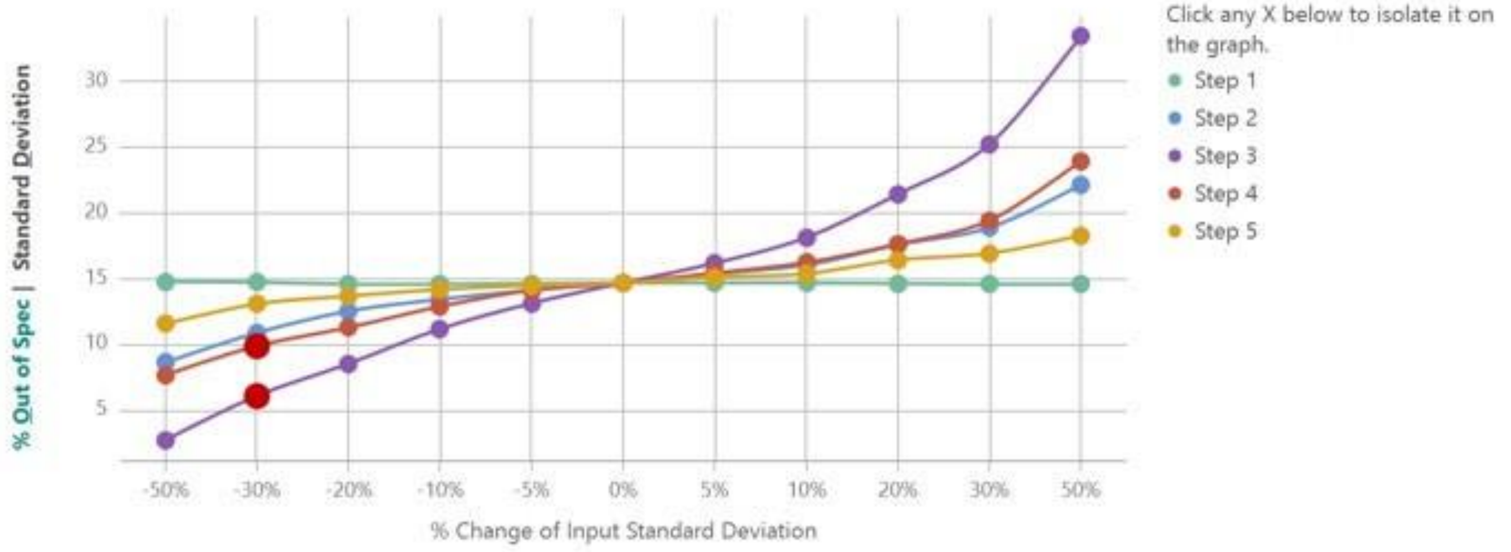
Summary Statistics

N	50,000
Mean	25.6707
Standard Deviation	5.5192

Sensitivity Analysis 2nd Choice

Sensitivity Analysis for Total Cycle Time

Use the graph to understand how changing the standard deviation of an input (X) affects the output (Y). Look for inputs that may improve the capability (sloped lines) and inputs that may allow less stringent tolerances (flat lines).



To change an input setting, click a point on the graph or choose a % change value from the table below.

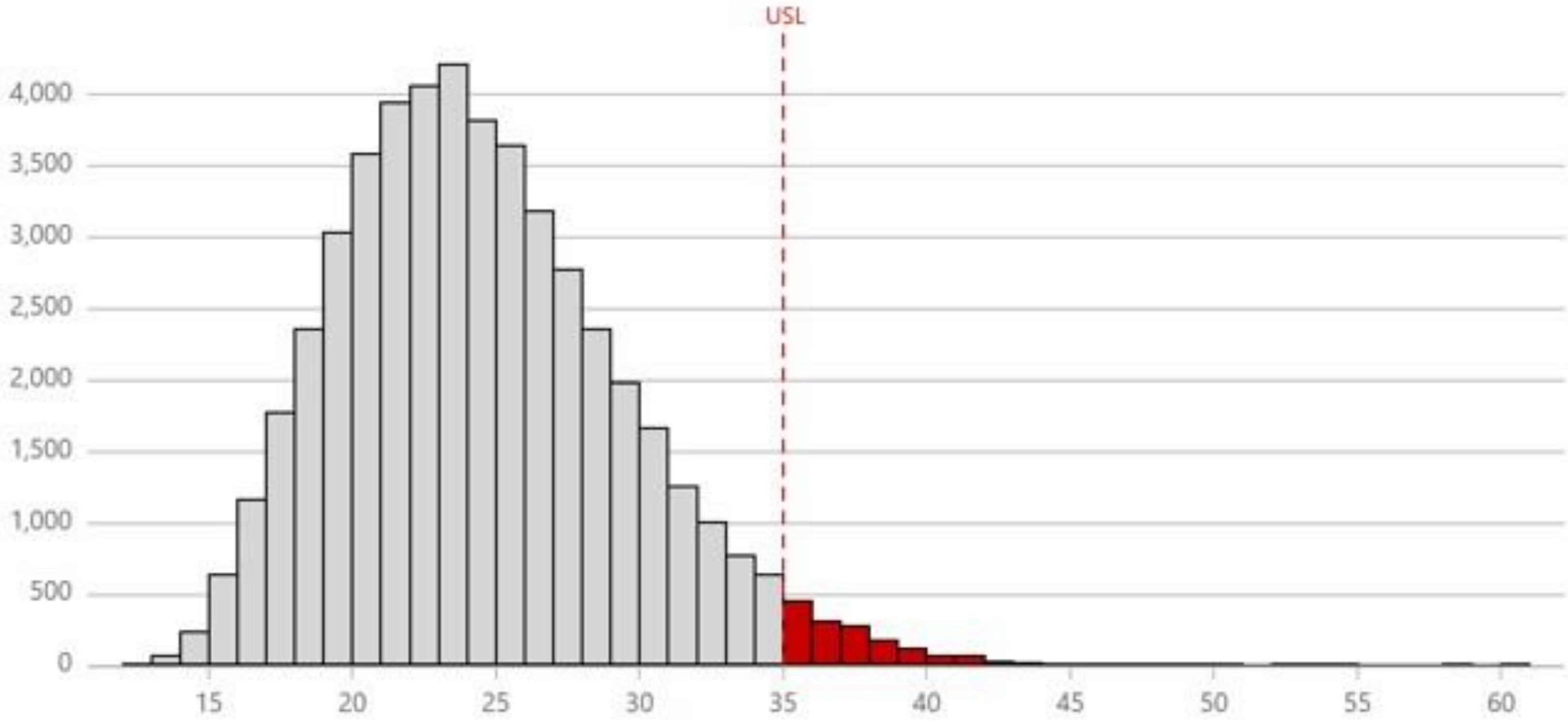
Inputs (X)	Current St Dev	% Change	Proposed St Dev
Step 1	0.57735	0%	—
Step 2	2.7463	0%	—
Step 3	4.592	-30%	3.2144
Step 4	3.1489	-30%	2.2043
Step 5	1.5156	0%	—

Reduce Steps 3 and 4 by 30%

Sensitivity Analysis 2nd Results

Sensitivity Analysis Results

Total Cycle Time



Using the new standard deviation values, the simulation indicates that you can expect 3.29% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.5617. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.5617

% Out of Spec

3.29%

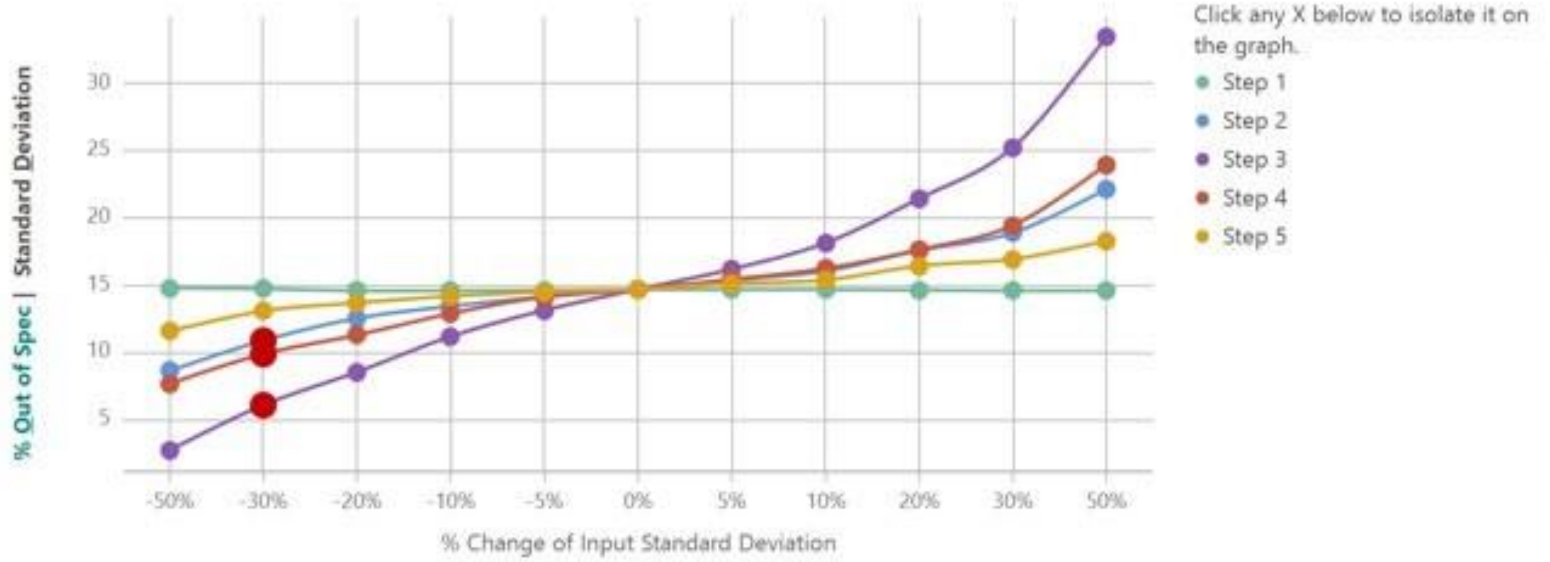
Summary Statistics

N	50,000
Mean	24.5221
Standard Deviation	5.0429

Sensitivity Analysis 3rd Choice

Sensitivity Analysis for Total Cycle Time

Use the graph to understand how changing the standard deviation of an input (X) affects the output (Y). Look for inputs that may improve the capability (sloped lines) and inputs that may allow less stringent tolerances (flat lines).



To change an input setting, click a point on the graph or choose a % change value from the table below.

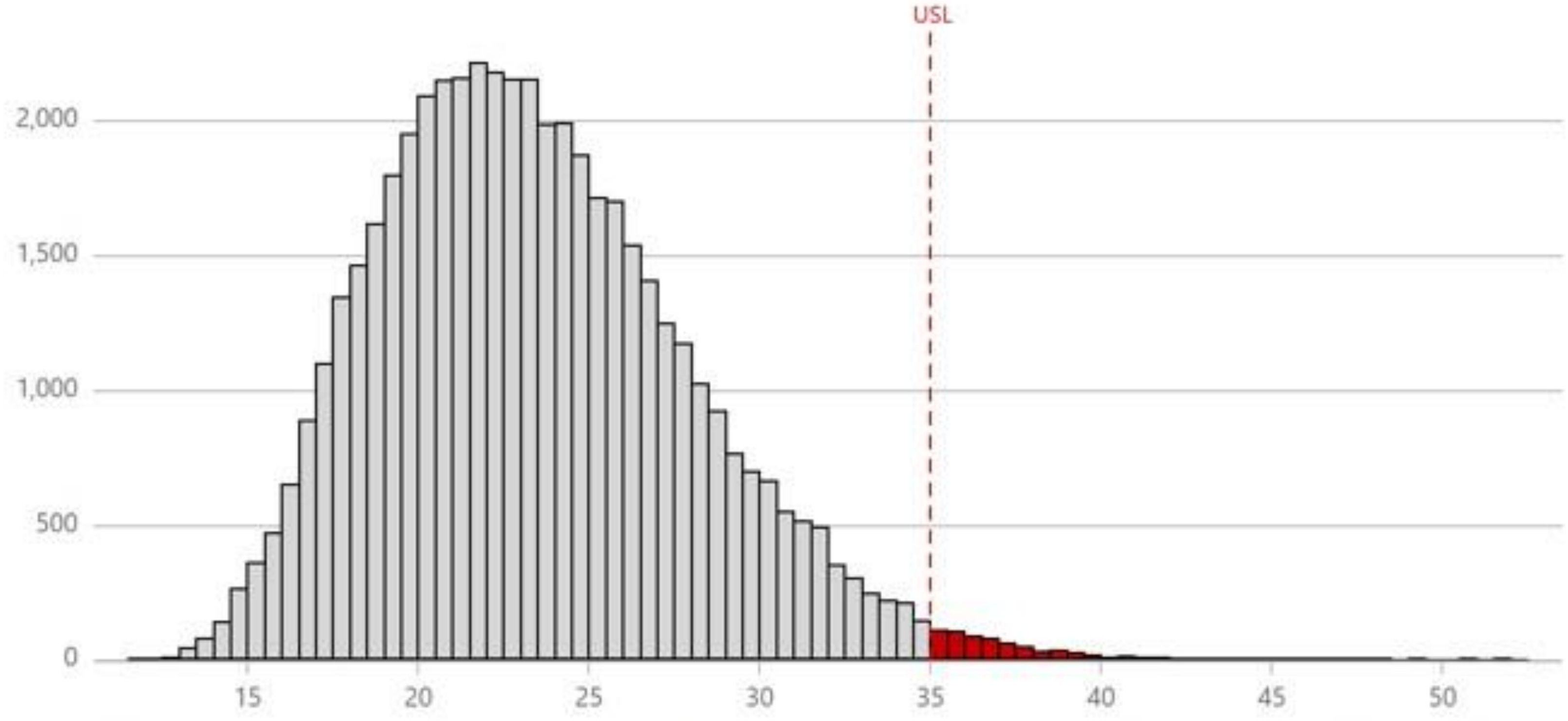
Inputs (X)	Current St Dev	% Change	Proposed St Dev
Step 1	0.57735	0%	—
Step 2	2.7463	-30%	1.9224
Step 3	4.592	-30%	3.2144
Step 4	3.1489	-30%	2.2043
Step 5	1.5156	0%	—

Reduce Steps 2, 3 and 4 by 30%

Sensitivity Analysis 3rd Results

Sensitivity Analysis Results

Total Cycle Time



Using the new standard deviation values, the simulation indicates that you can expect 1.52% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.6432. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.6432

% Out of Spec

1.52%

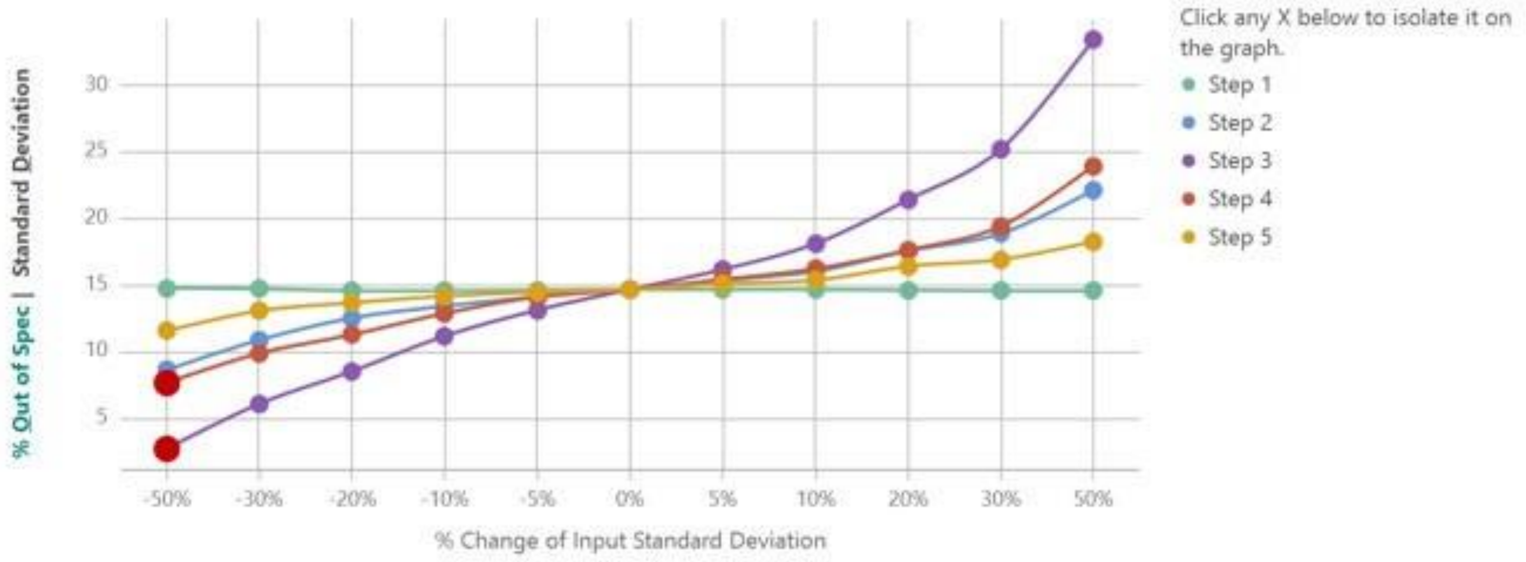
Summary Statistics

N	50,000
Mean	23.427
Standard Deviation	4.6355

Sensitivity Analysis 4th Choice

Sensitivity Analysis for Total Cycle Time

Use the graph to understand how changing the standard deviation of an input (X) affects the output (Y). Look for inputs that may improve the capability (sloped lines) and inputs that may allow less stringent tolerances (flat lines).



To change an input setting, click a point on the graph or choose a % change value from the table below.

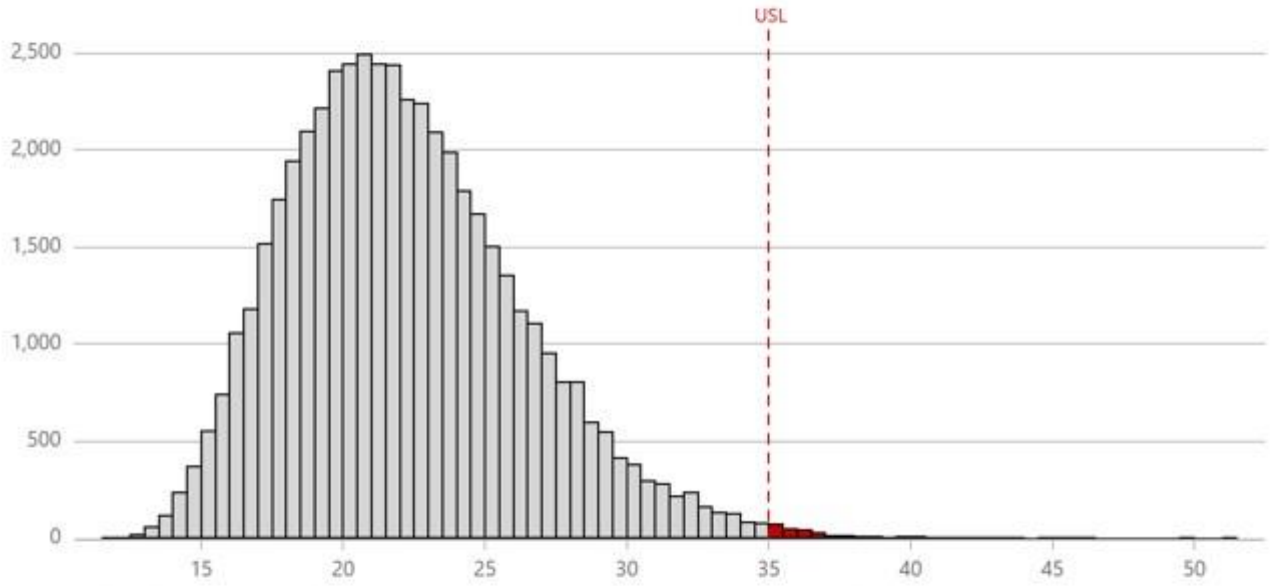
Inputs (X)	Current St Dev	% Change	Proposed St Dev
Step 1	0.57735	0%	—
Step 2	2.7463	0%	—
Step 3	4.592	-50%	2.296
Step 4	3.1489	-50%	1.5745
Step 5	1.5156	0%	—

Reduce Steps 3 and 4 by 50%

Sensitivity Analysis 4th Results

Sensitivity Analysis Results

Total Cycle Time



Using the new standard deviation values, the simulation indicates that you can expect 0.73% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.7376. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.7376

% Out of Spec

0.73%

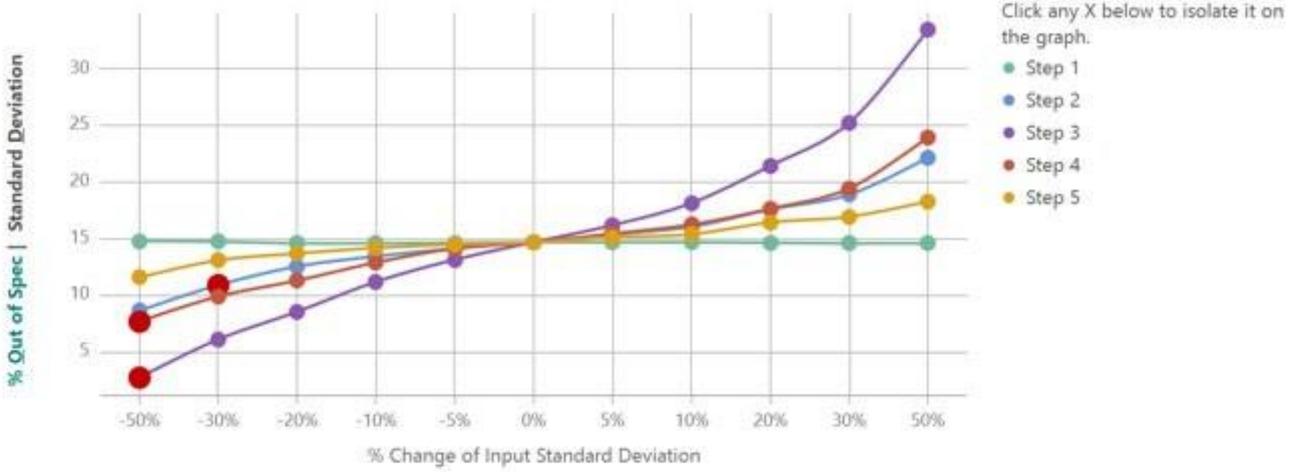
Summary Statistics

N	50,000
Mean	22.2361
Standard Deviation	4.2441

Sensitivity Analysis 5th Choice

Sensitivity Analysis for Total Cycle Time

Use the graph to understand how changing the standard deviation of an input (X) affects the output (Y). Look for inputs that may improve the capability (sloped lines) and inputs that may allow less stringent tolerances (flat lines).



To change an input setting, click a point on the graph or choose a % change value from the table below.

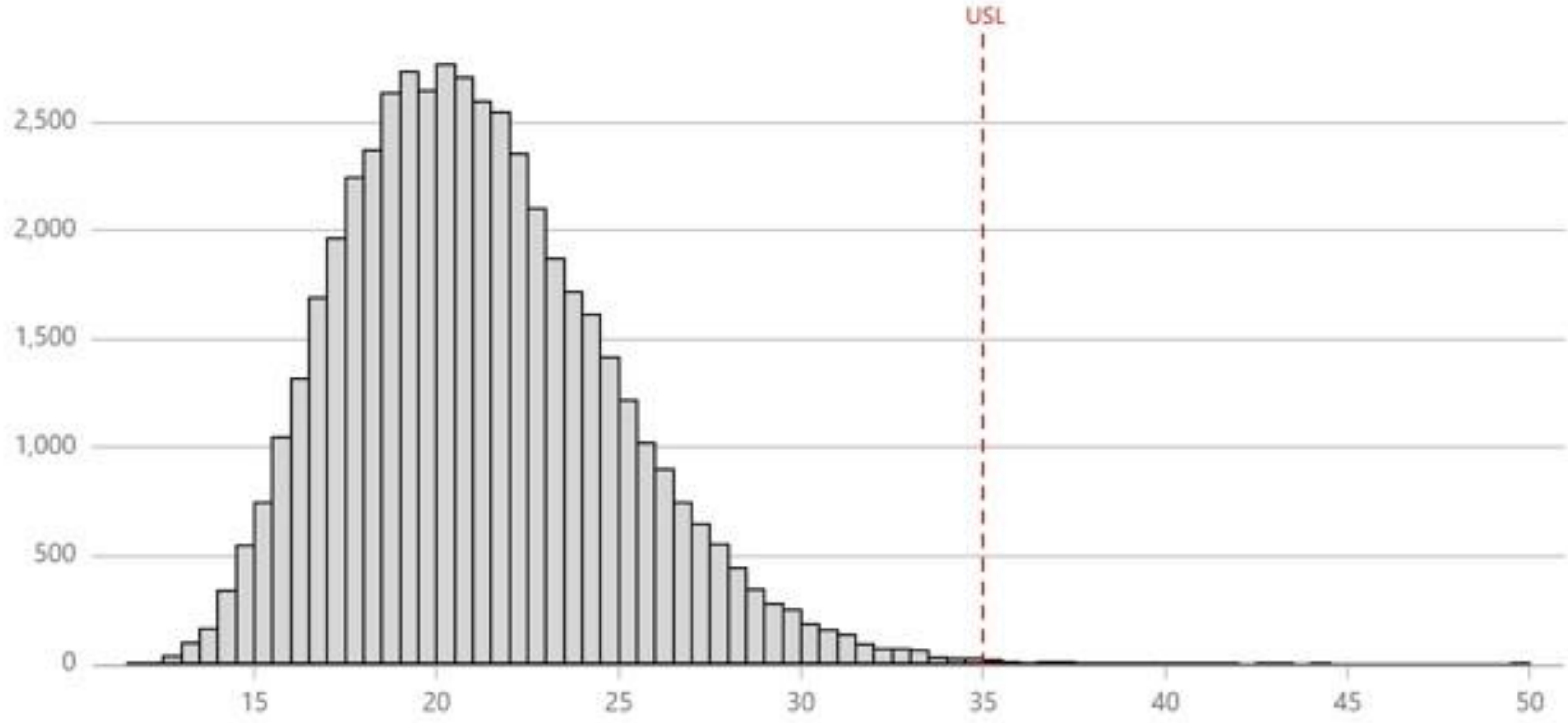
Inputs (X)	Current St Dev	% Change	Proposed St Dev
Step 1	0.57735	0%	—
Step 2	2.7463	-30%	1.9224
Step 3	4.592	-50%	2.296
Step 4	3.1489	-50%	1.5745
Step 5	1.5156	0%	—

Reduce Step 2 by 30%, Steps 3 and 4 by 50%

Sensitivity Analysis 5th Results

Sensitivity Analysis Results

Total Cycle Time



Process Performance (Cpk)

0.9335

% Out of Spec

0.22%

Summary Statistics

N	50,000
Mean	21.1656
Standard Deviation	3.756

Using the new standard deviation values, the simulation indicates that you can expect 0.22% of the *Total Cycle Time* values to fall outside of the specification limits. This corresponds to a Cpk of 0.9335. A generally accepted minimum value of Cpk is 1.33.

Synopsis of Analyses Using Observed Data

ANALYSIS	Cpk	% Out of Spec
Initial Sim	0.274	16.69%
Parameter Opt.	0.311	13.66%
Sensitivity #1	0.4438	5.97%
Sensitivity #2	0.5617	3.29%
Sensitivity #3	0.6432	1.52%
Sensitivity #4	0.7376	0.73%
Sensitivity #5	0.9335	0.22%

Take-aways

- Through this session, you should have:
  - Become familiar with Monte Carlo simulation, how to run a simulation to get initial results and possible next steps in analysis
  - An understanding of how to use “what if” scenarios and vary input parameters to see what the results will be
  - Learned how to do a sensitivity analysis to target improvements to the step that will gain you the greatest overall improvement

Questions?

Thoughts?

Comments?

Contact information:

Charles Cox | Chuck@TheOpExExpert.com

Free innovation and operational excellence presentations, publications, and insights: www.TheOpExExpert.com

Free book chapter: innovatingleansixsigma.com

ASQ

Excellence Through Quality®