AREVA

forward-looking energy

FIT

FUELS INTEGRATED TRAINING

Module 2

Basic Statistics Refresher

AREVA
forward-looking energy

Outline

- INTRODUCTION
- NORMAL DISTRIBUTION
- TOLERANCE INTERVALS
- RESPONSE SURFACE MODELS

Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- Normal Distribution
  - Explain why “2-sigma” is commonly used to describe 95% coverage of a normal distribution
- Tolerance Intervals
  - Describe a 95/95 tolerance interval
  - Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval
- Response Surface Models
  - Define a Response Surface Model

Outline

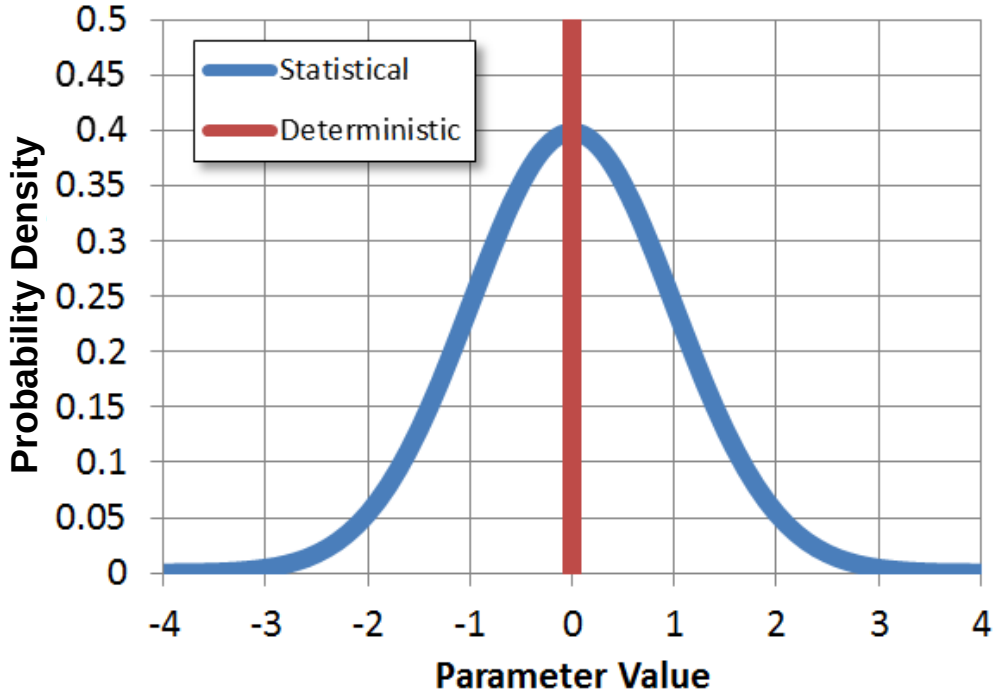
- INTRODUCTION
- NORMAL DISTRIBUTION
- TOLERANCE INTERVALS
- RESPONSE SURFACE MODELS

Learning Objectives

- **Introduction**
  - **Describe the difference between a bias and a random uncertainty**
- **Normal Distribution**
  - **Explain why "2-sigma" is commonly used to describe 95% coverage of a normal distribution**
- **Tolerance Intervals**
  - **Describe a 95/95 tolerance interval**
  - **Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval**
- **Response Surface Models**
  - **Define a Response Surface Model**

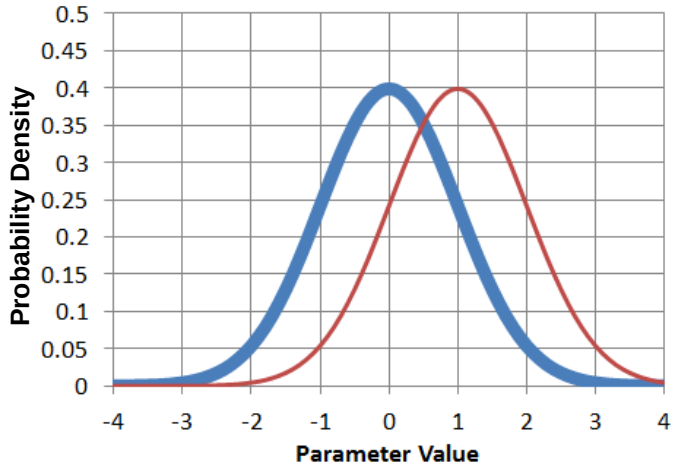
Deterministic vs. Statistical

- **Statistics allows for the treatment of random variability.**
  - **Deterministic – Single value used to represent a Parameter.**
  - **Statistical – Distribution of value used to represent a parameter including random variability.**

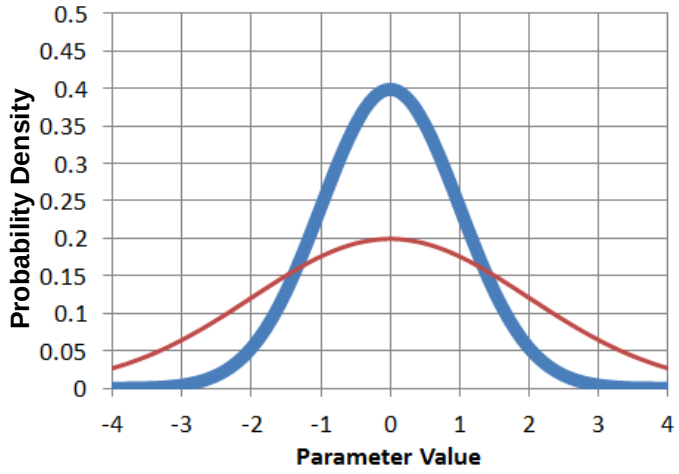


Types of Uncertainty

- **Bias – Shifts the mean of the distribution to the left or to the right**



- **Random – Impacts the shape of the distribution**



Types of Uncertainty

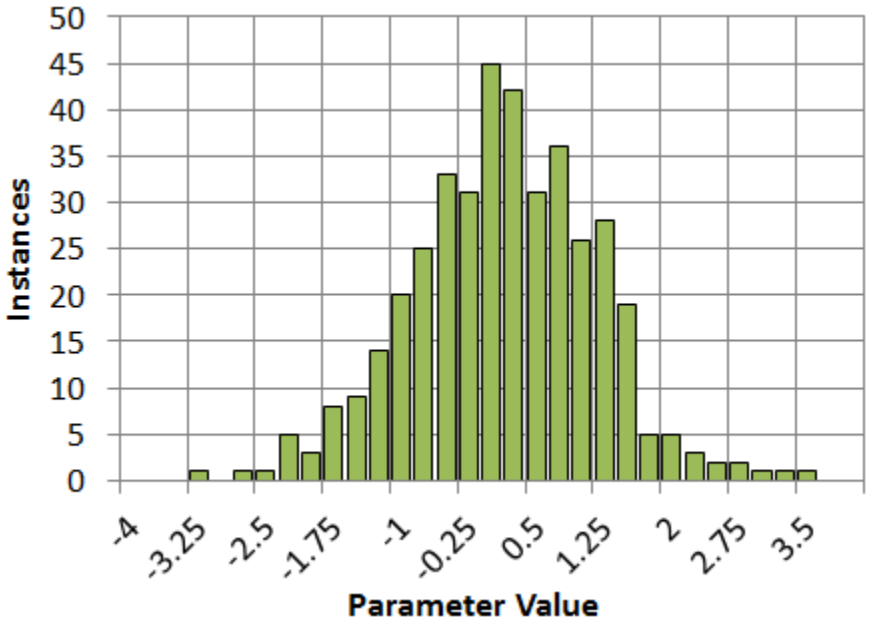
- **A base assumption throughout the setpoint methodology is that uncertainty parameters may be treated as symmetric and normally distributed.** Not always the case in reality

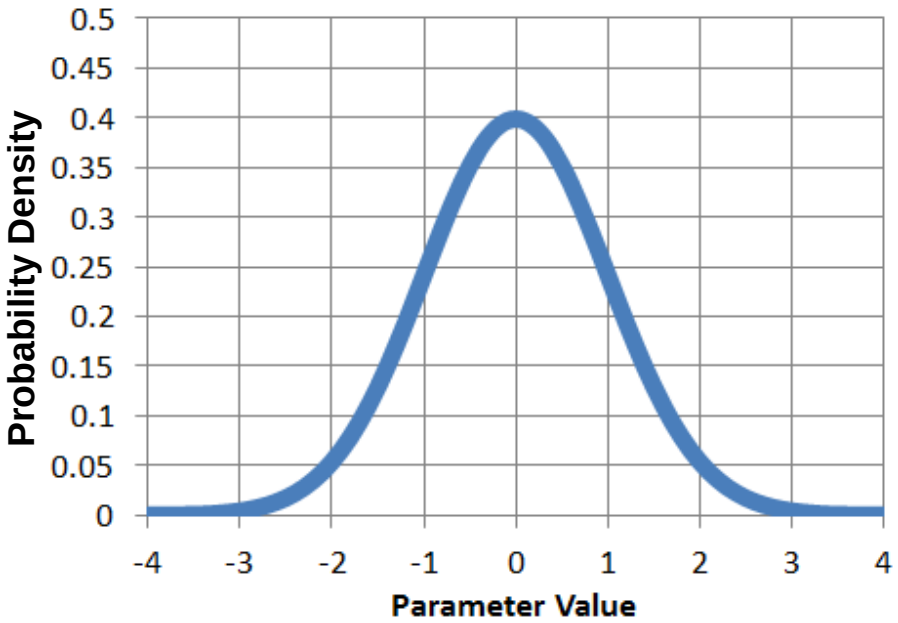
Example T_{avg} uncertainty = +4.0F/-4.8F	Bias	= -0.4F
	Random	= +-4.4F
Option 1 ◆ Apply bounding random uncertainty	Bias	= -0.0F
	Random	= +-4.8F
Option 2 ◆ Apply bias to nominal setting	Bias	= -0.4F
	Random	= +-4.4F

- **Pay extra attention in these situations**

Visualization

- A Histogram is used to visualize how a collection of data points is distributed.
- Probability Density Functions (PDF) are often used to analytically estimate statistical distributions.





Review Learning Objectives

- **Introduction**
  - **Describe the difference between a bias and a random uncertainty**
- **Normal Distribution**
  - **Explain why "2-sigma" is commonly used to describe 95% coverage of a normal distribution**
- **Tolerance Intervals**
  - **Describe a 95/95 tolerance interval**
  - **Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval**
- **Response Surface Models**
  - **Define a Response Surface Model**

Outline

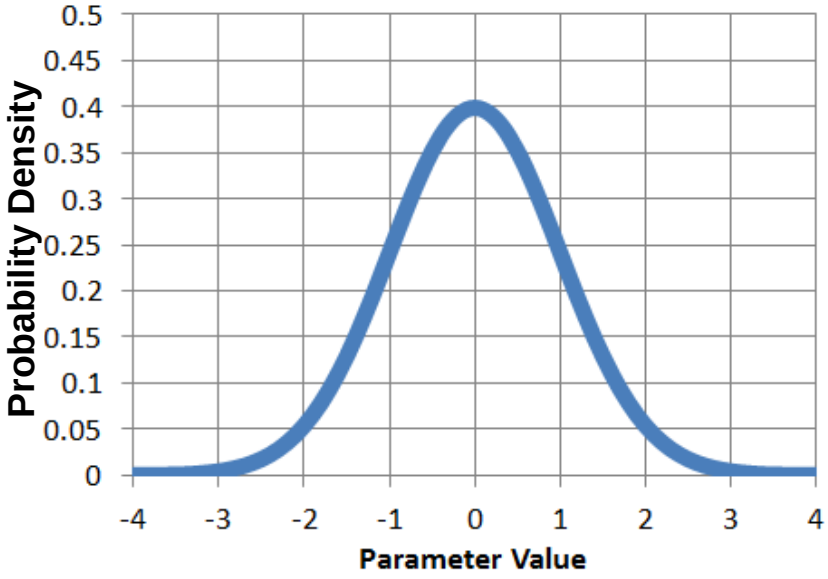
- INTRODUCTION
- NORMAL DISTRIBUTION
- TOLERANCE INTERVALS
- RESPONSE SURFACE MODELS

Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- **Normal Distribution**
  - **Explain why "2-sigma" is commonly used to describe 95% coverage of a normal distribution**
- Tolerance Intervals
  - Describe a 95/95 tolerance interval
  - Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval
- Response Surface Models
  - Define a Response Surface Model

Normal Distribution

- **Many of the setpoint calculations use probability distributions to model the real world variability in input parameters.**
- **Simplified models are typically used.**
  - **Most common is a Normal distribution.**



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

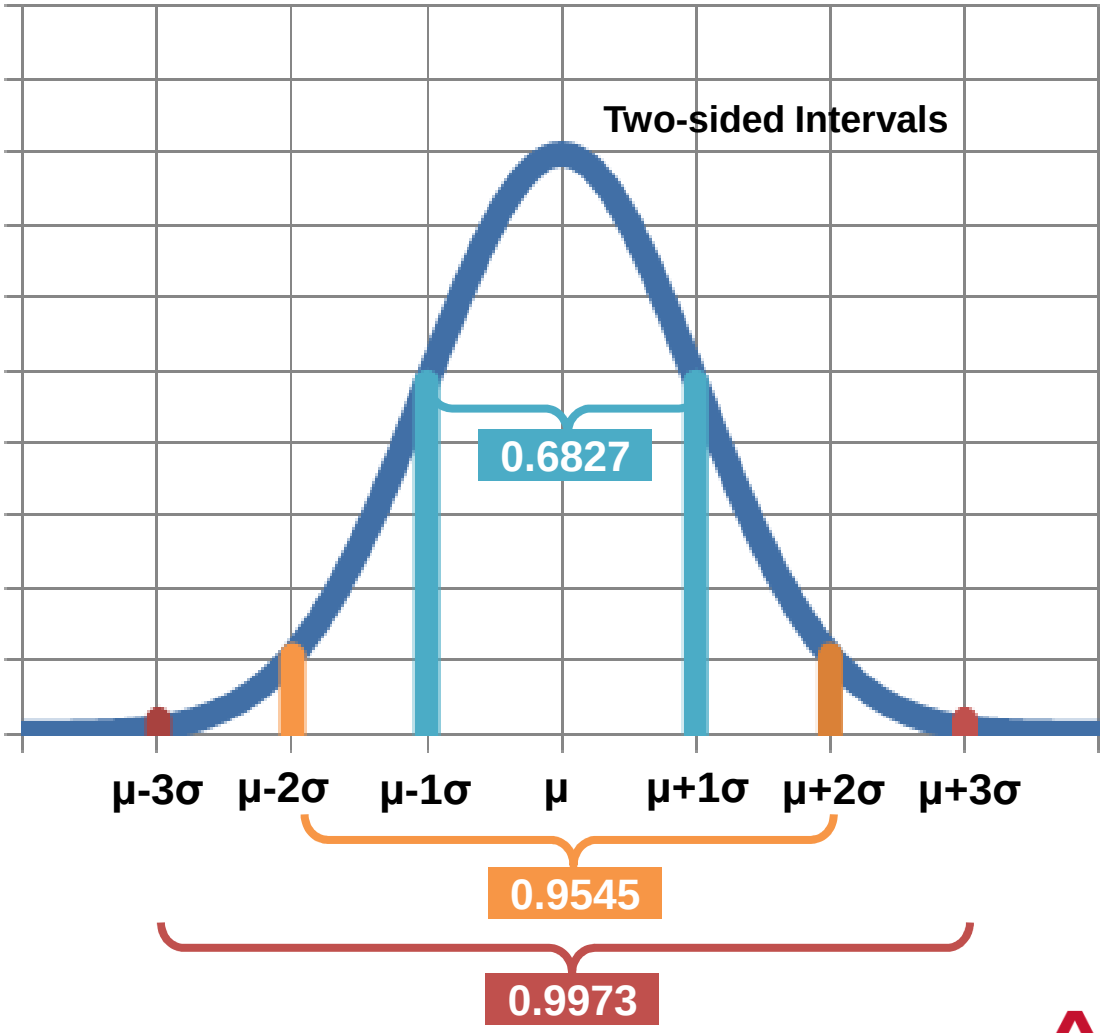
Where,

μ = mean

σ = standard deviation

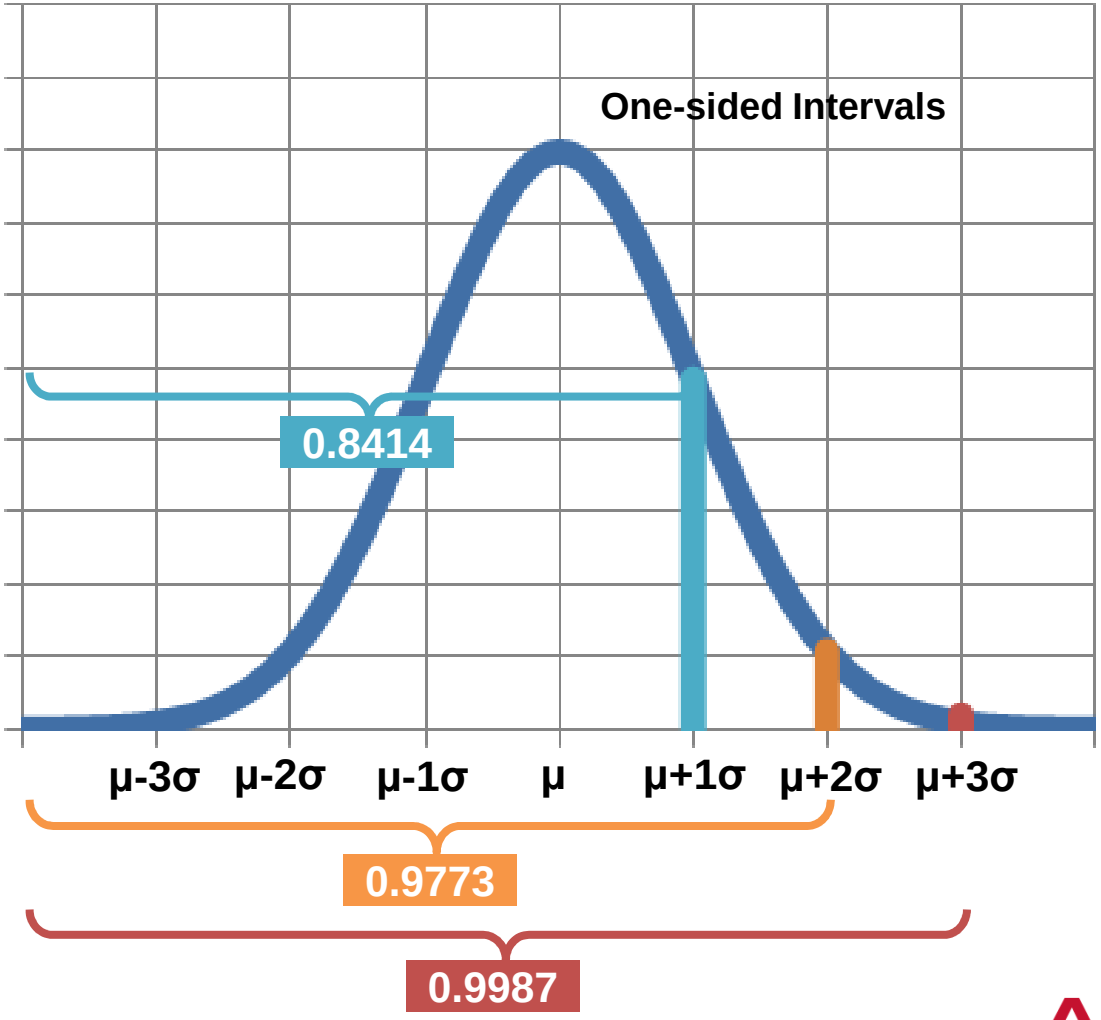
Normal Distribution

- **A normal curve represents probability as the area under the curve.**
  - **Typically calculated using tables because there is not a closed form solution for the integral.**
  - **Student's T distribution**
  - **Note that the 2-sigma interval covers just over 95% of the data.**



Normal Distribution

- **A normal curve represents probability as the area under the curve.**
  - **Typically calculated using tables because the is not closed form solution for the integral.**



Review of Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- **Normal Distribution**
  - **Explain why "2-sigma" is commonly used to describe 95% coverage of a normal distribution**
- Tolerance Intervals
  - Describe a 95/95 tolerance interval
  - Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval
- Response Surface Models
  - Define a Response Surface Model

Outline

- INTRODUCTION
- NORMAL DISTRIBUTION
- TOLERANCE INTERVALS
- RESPONSE SURFACE MODELS

Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- Normal Distribution
  - Explain why “2-sigma” is commonly used to describe 95% coverage of a normal distribution
- **Tolerance Intervals**
  - **Describe a 95/95 tolerance interval**
  - **Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval**
- Response Surface Models
  - Define a Response Surface Model

Tolerance Interval

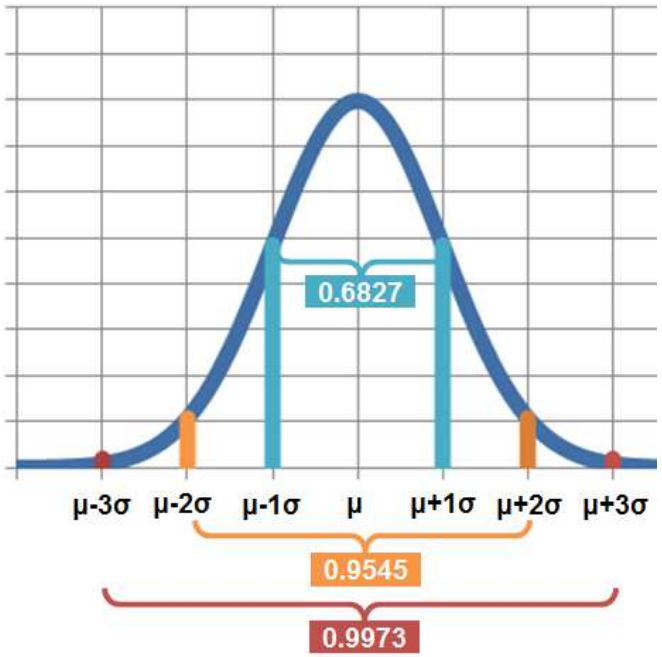
- A finite data set is inadequate to characterize a parameter's variability with 100% certainty.
- A tolerance interval has two numbers associated with it, namely a confidence level and a coverage level.
- The interval is made so that we can have specified confidence that at least the specified portion of the entire population is covered by the interval.
- For example, a 99/95 tolerance interval means that there is 99% confidence that 95% of the population will be covered by the given interval.
  - The order of the terms is not always treated consistently (i.e. 99/95 vs. 95/99).
  - Luckily, we usually use 95/95 so it does not matter.

Tolerance Interval

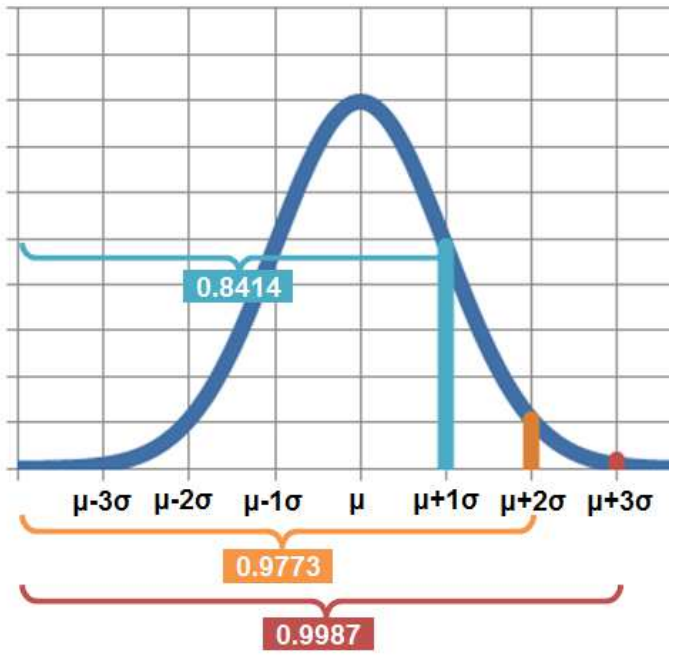
- **For example, a 99/95 tolerance interval means that there is 99% confidence that 95% of the population will be covered by the given interval.**
- **The 99% refers to the confidence of the interval.**
- **This is typically impacted by the number of samples available.**
  - **A large uncertainty factor must be applied if a small number of samples is available or if a high confidence is desired.**
  - **A small uncertainty factor can be applied is a large number of samples is available or if a low confidence is desired.**
  - **The confidence of a best estimate calculation is 50%.**

Tolerance Interval

- **The 95% is the portion of the population being covered.**
- **Coverage can be one-sided or two-sided.**



two-sided



one-sided

Tolerance Interval

- **Tolerance intervals for normal distributions are of the form:**

$\bar{X} \pm k * s$ **(two−sided)** $\bar{X} - k * s$ **(lower)** $\bar{X} + k * s$ **(upper)**

- **The constants k are typically referred to as “k-factors” and are tabulated.**
- **The K factor for a two-sided 95/95 tolerance interval with infinite samples is 1.96**
- **The K factor for a one-sided 95/95 tolerance interval with infinite samples is 1.645**

$\bar{X}$ **(is the sample population mean)** s **(is the sample population standard deviation)**

Review of Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- Normal Distribution
  - Explain why “2-sigma” is commonly used to describe 95% coverage of a normal distribution
- **Tolerance Intervals**
  - **Describe a 95/95 tolerance interval**
  - **Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval**
- Response Surface Models
  - Define a Response Surface Model

Outline

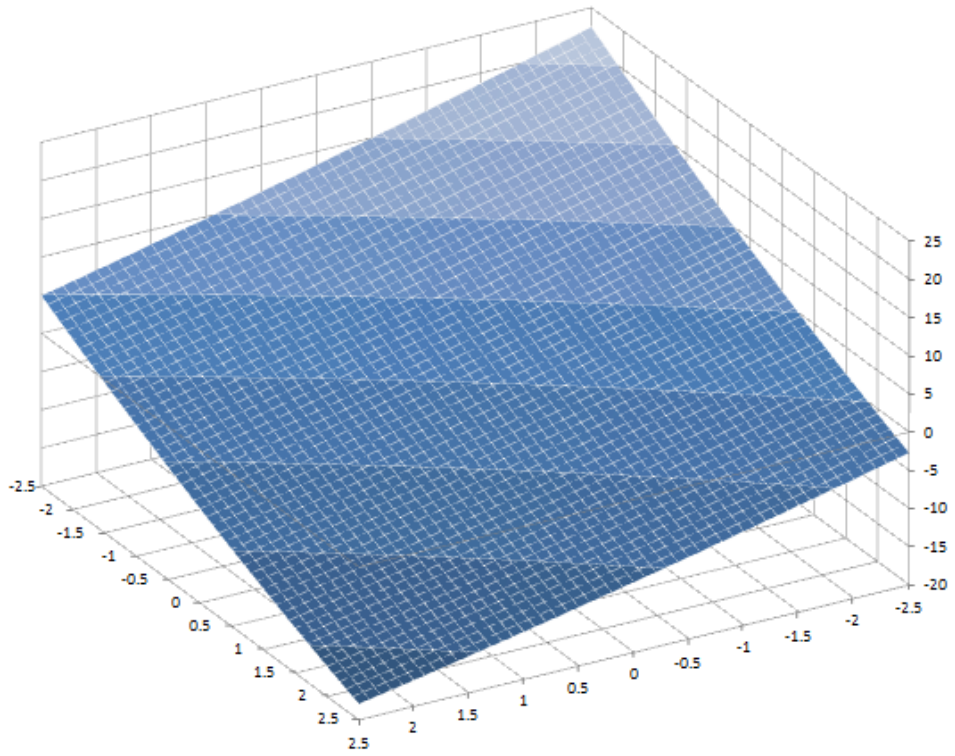
- INTRODUCTION
- NORMAL DISTRIBUTION
- TOLERANCE INTERVALS
- **RESPONSE SURFACE MODELS**

Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- Normal Distribution
  - Explain why “2-sigma” is commonly used to describe 95% coverage of a normal distribution
- Tolerance Intervals
  - Describe a 95/95 tolerance interval
  - Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval
- **Response Surface Models**
  - **Define a Response Surface Model**

Response Surface Model (RSM)

- A response surface is a multi-dimensional fit of a particular response to a set of input parameters.
- Typically used to estimate complex phenomenon in an efficient way.



AREVA forward-looking energy

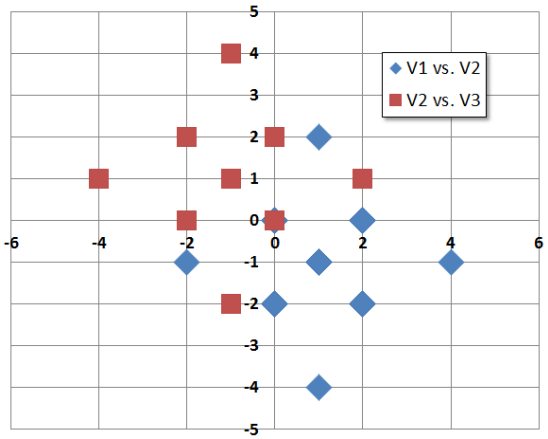
Response Surface Model

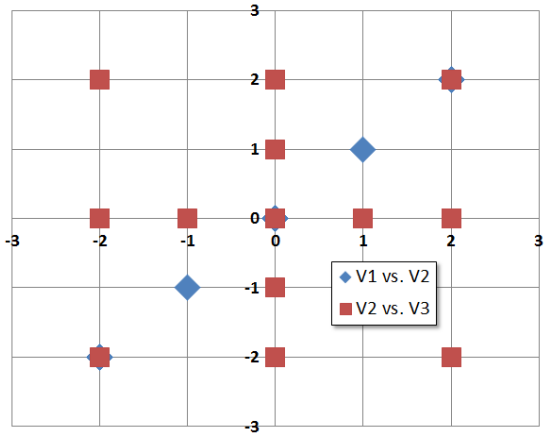
- **To develop an RSM, a set of experimental results characterizing the design space is required.**
  - **For DNBR calculation, the experiments are explicit XCOBRA-IIIC runs**
  - **The design space is defined by the min and max of each input parameter**
  - **Typically inputs are varied at integer multiples of their standard deviation**
- **One option is to evaluate all possible combinations of each parameter at a given set of levels.**
  - **For example, assume each parameter can be at -2σ, -1σ, 0σ, 1σ, or 2σ**
  - **Running all combinations of 10 parameters at 5 levels results in 9,765,625 XCOBRA-IIIC runs**

Response Surface Model

Design of Experiments (DOE)

- **A design of experiments can be used to limit the number of explicit runs needed to build an RSM while minimizing the loss of information**
- **Many types of DOE exist**
  - Box-Benhken design
  - Plackett-Burman
  - Cubic centered design



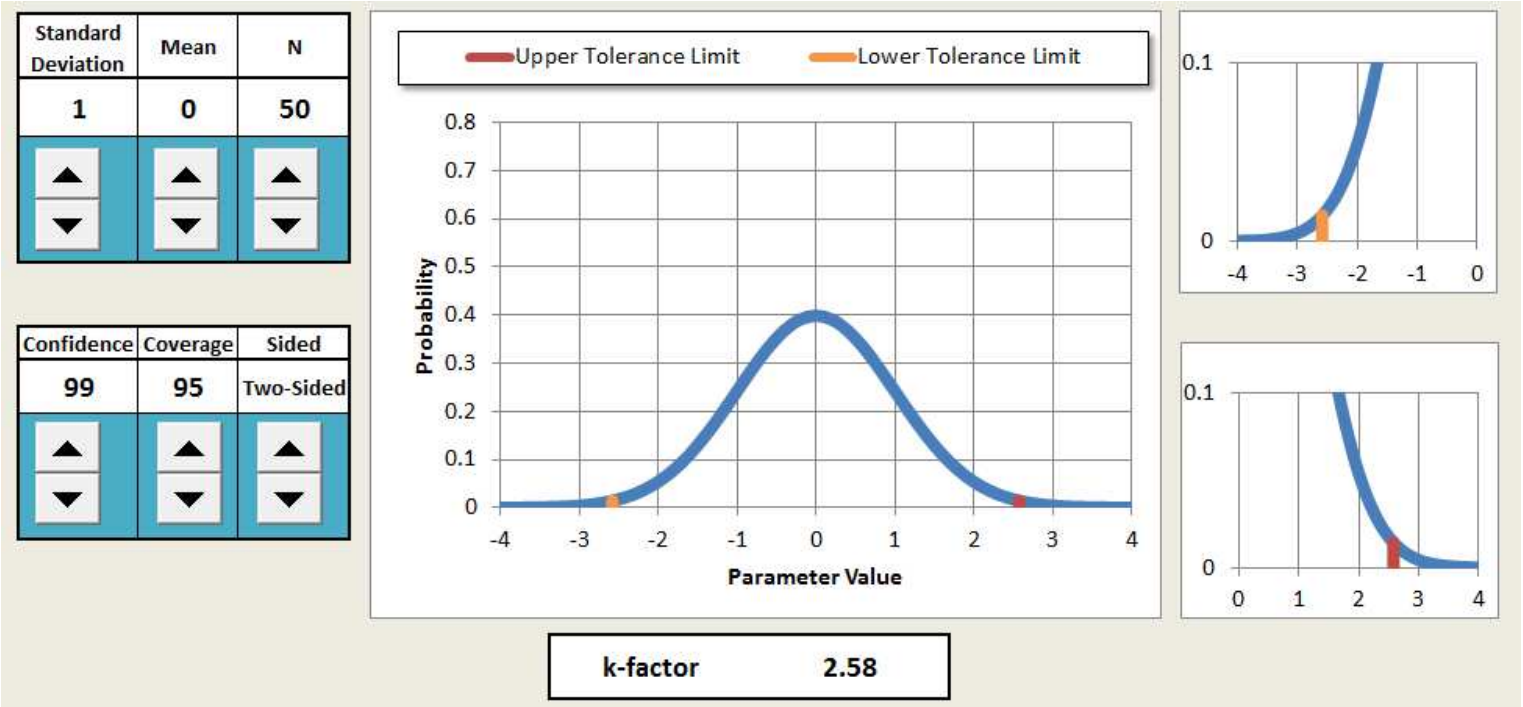


Review of Learning Objectives

- Introduction
  - Describe the difference between a bias and a random uncertainty
- Normal Distribution
  - Explain why “2-sigma” is commonly used to describe 95% coverage of a normal distribution
- Tolerance Intervals
  - Describe a 95/95 tolerance interval
  - Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval
- Response Surface Models
  - **Define a Response Surface Model**

Exercise 2.1

- **Companion Notebook > Statistics Tab > Exercise 2.1**
  - **Tolerance Intervals**



Review of Learning Objectives

- **Introduction**
  - Describe the difference between a bias and a random uncertainty
- **Normal Distribution**
  - Explain why “2-sigma” is commonly used to describe 95% coverage of a normal distribution
- **Tolerance Intervals**
  - Describe a 95/95 tolerance interval
  - Explain the difference between a one-sided tolerance interval and a two-sided tolerance interval
- **Response Surface Models**
  - Define a Response Surface Model