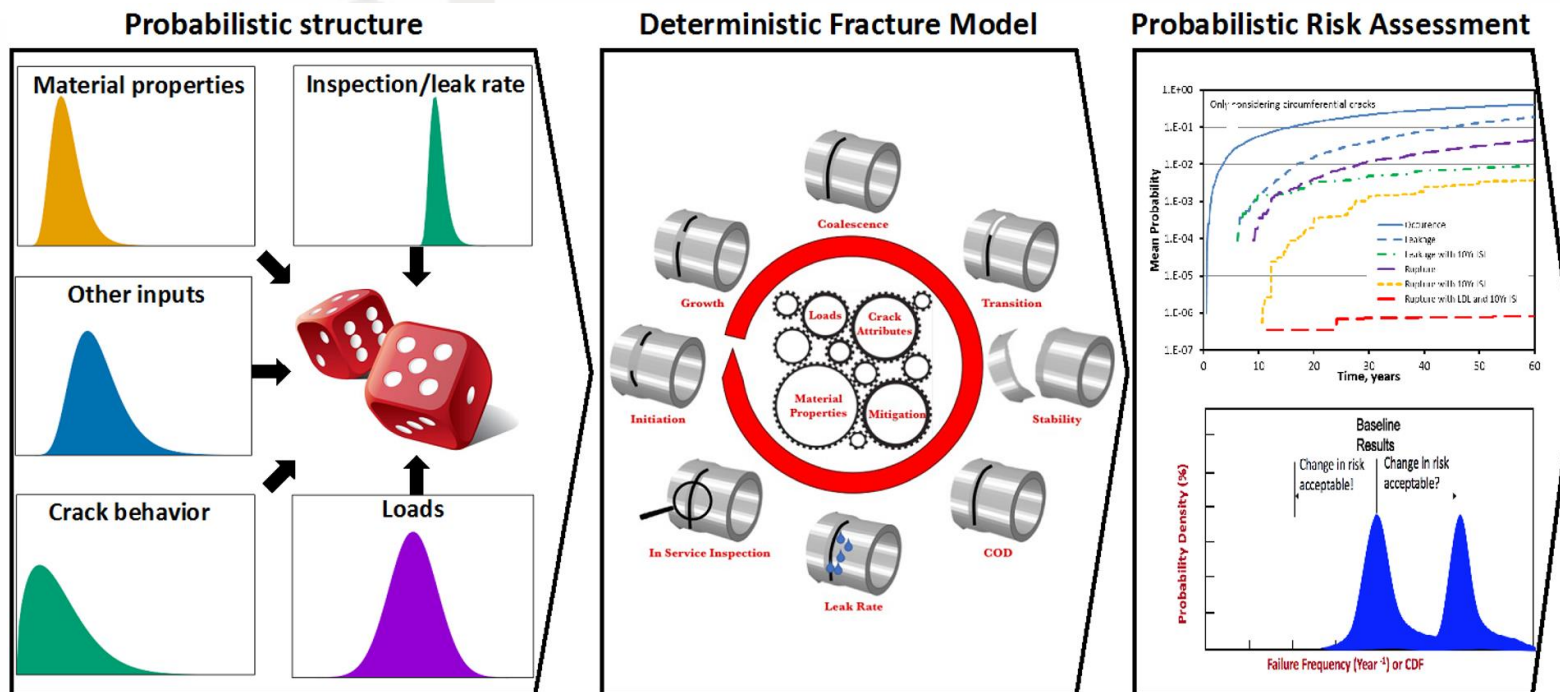




MODULE 1: PFM BACKGROUND



Email: xlpr_team@sandia.gov

Site: <https://connect.sandia.gov/sites/xLPR/SitePages/homepage.aspx>



- **At the end of this module, students will be able to:**
 - Understand and compare deterministic and probabilistic fracture mechanics approaches
 - Understand the fundamental building blocks of a probabilistic fracture mechanics (PFM) analysis including:
 - Characterization of uncertainty
 - Separation of uncertainty
 - Methods for sampling inputs for Monte Carlo analysis
 - Interpretation of PFM results



- Probabilistic Fracture Mechanics (PFM) Background
- Comparing deterministic and probabilistic approaches
- Components of a PFM analysis
 - Characterizing input uncertainty
 - Separation of uncertainty: Aleatory vs. Epistemic
 - Sampling structure and sampling schemes
 - Understanding PFM results
- Example of a PFM analysis
 - Basic mechanics example
 - Example of an xLPR V2.0 PFM analysis

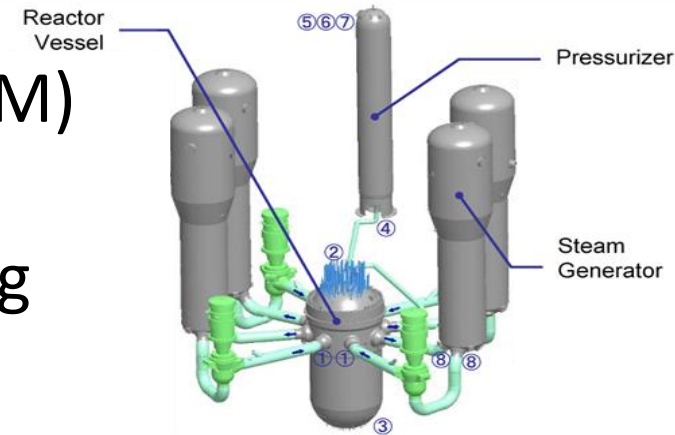


Probabilistic Fracture Mechanics (PFM) Background

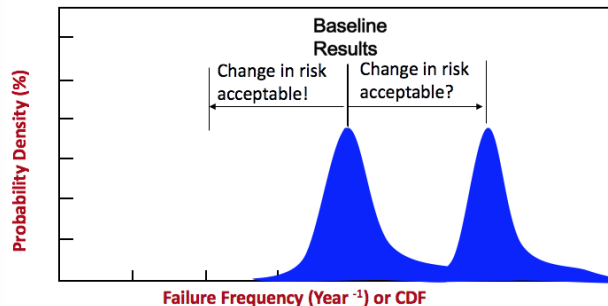


MOTIVATION BEHIND CONDUCTING A PFM ANALYSIS

- **Probabilistic Fracture Mechanics (PFM)** has been increasingly used in the fitness for service assessment of aging piping, reactor vessels, and steam generator tubing in recent years

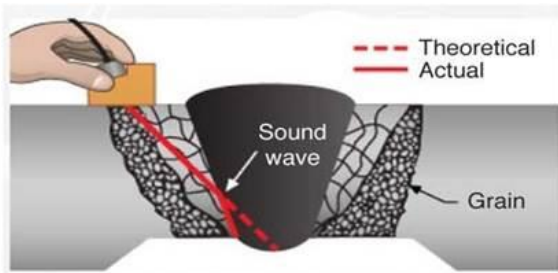


Hirano, EJAM (2010)



- **Used in:**

- **Non-Destructive Examination (NDE)**
- **Risk Informed In-Service Inspection (ISI)**
- Enhance the technical basis of **Probabilistic Risk Assessment (PRA)**





WHY DO WE PERFORM A PROBABILISTIC ANALYSIS?

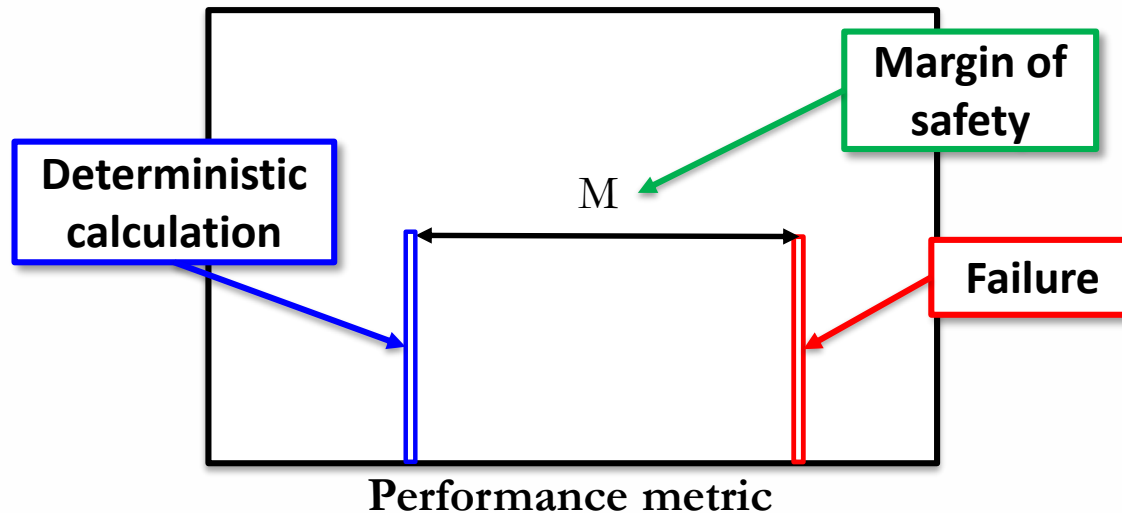
- To better understand **performance margins and uncertainties**
- Many applications have **probabilistic performance requirements**:
 - Probability of an undesirable event happening $< 10^{-m}$
- To provide **consistent set of criteria on systems** so that resources can be focused where needed most
- **Qualification support**:
 - Level of confidence in design
 - Body of evidence that the system meets its design requirements



Comparing deterministic and probabilistic approaches



MARGIN OF SAFETY: DETERMINISTIC APPROACH



- **Deterministic approach** assumes all significant parameters defining the problem are known
- Where uncertainties exist (e.g., materials properties) **conservative bounding values** are assumed
- Safety factors are imposed to ensure satisfactory margins against uncertainties



- Deterministic approach uses fixed inputs that generate a single output

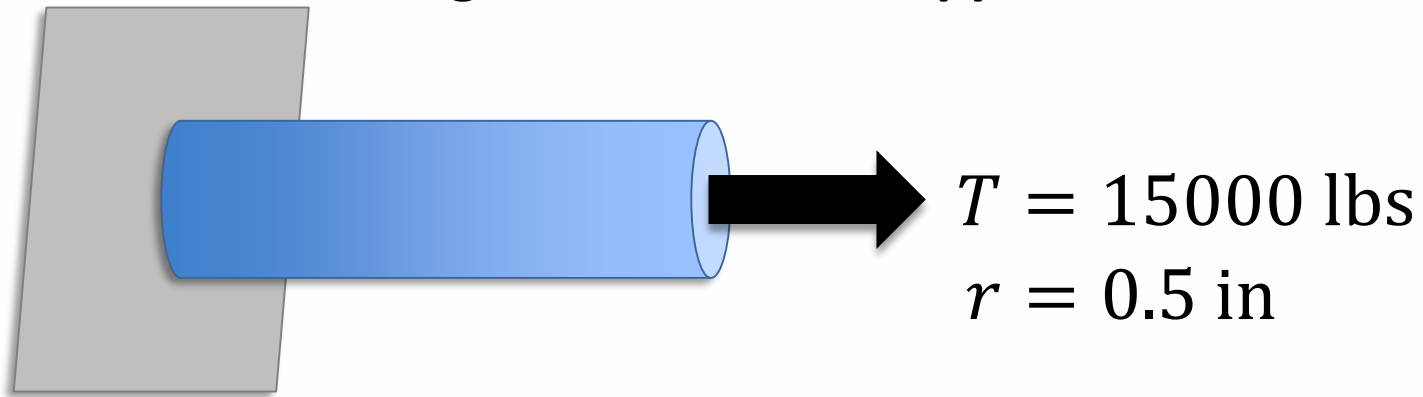


- For example:
 X = Weld Residual Stress (WRS) profile across weld thickness
 Y = Occurrence of a through-wall crack



EXERCISE: DETERMINISTIC AXIAL STRESS IN A BAR

Example Problem: Calculate the stress in an axially loaded bar and determine if it will fail using a deterministic approach



- Flow stress: $\sigma_f = 55.5 \text{ Ksi}$
- Safety factor: $SF = 2.77$ (ASME Section III Code Allowable Stress; SA-36)
- Primary membrane stress: $\sigma_m = \frac{\text{Load}}{\text{Area}} = \frac{T}{\pi r^2}$

- Navigate to \Exercises\Module 1 and open the file “Module_1_Axial_Stress_in_a_Bar.xlsx”
- On the first tab, “Deterministic problem”, enter the parameter values
- Does the bar fail?

Deterministic example: Calculate the stress in an axially loaded bar and determine if it will fail

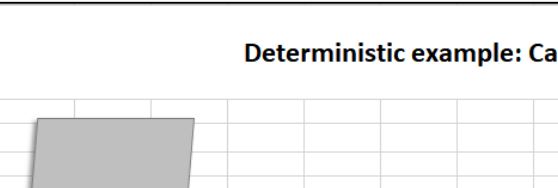


Diagram illustrating an axially loaded bar fixed to a wall on the left and pulled by a force T on the right.

Parameters:

- T [lbs]
- r [in]
- Flow Stress σ_f [Ksi]
- Safety Factor SF

Results:

- Primary Membrane Stress σ_m [Ksi]

Failure:

- Is $SF \times \sigma_m > \sigma_f$? [No = 0, Yes = 1 (Failure)]

Calculations:

$$T = 15000 \text{ lbs}$$
$$r = 0.5 \text{ in}$$
$$\sigma_m = \frac{\text{Load}}{\text{Area}} = \frac{T}{\pi r^2}$$
$$\sigma_f = 55.5 \text{ Ksi}$$
$$SF = 2.77$$



EXERCISE: DETERMINISTIC AXIAL STRESS IN A BAR

- **Solution:** Does the bar fail?
 - The bar does not fail, the primary membrane stress (19.099 Ksi) is less than the flow stress (55.5 Ksi)

Deterministic example: Calculate the stress in an axially loaded bar and determine if it will fail



Parameters				Results	Failure
T [lbs]	r [in]	Flow Stress σ_f [Ksi]	Safety Factor SF	Primary Membrane Stress σ_m [Ksi]	Is SF x $\sigma_m > \sigma_f$? [No = 0, Yes = 1 (Failure)]
15000	0.5	55.5	2.77	19.099	0

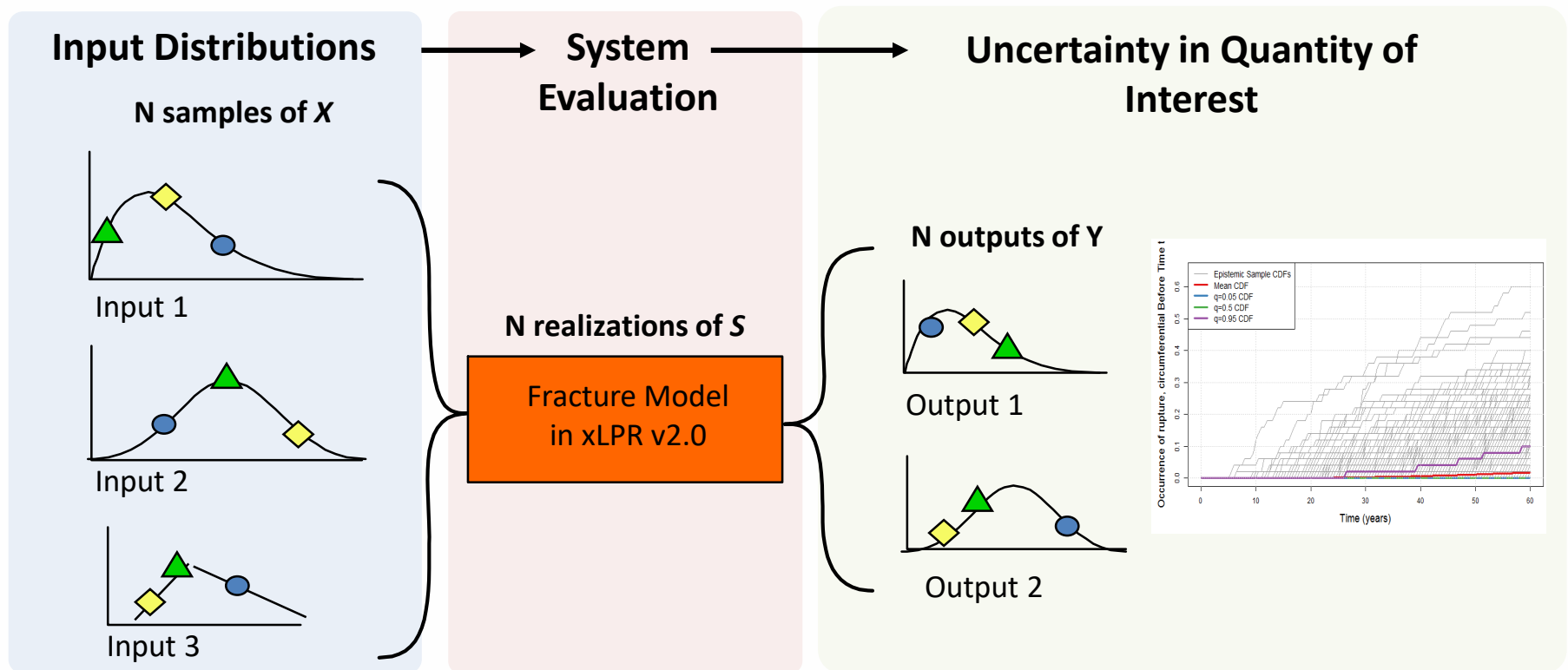
$$T = 15000 \text{ lbs}$$

$$r = 0.5 \text{ in}$$

$$\sigma_m = \frac{\text{Load}}{\text{Area}} = \frac{T}{\pi r^2}$$

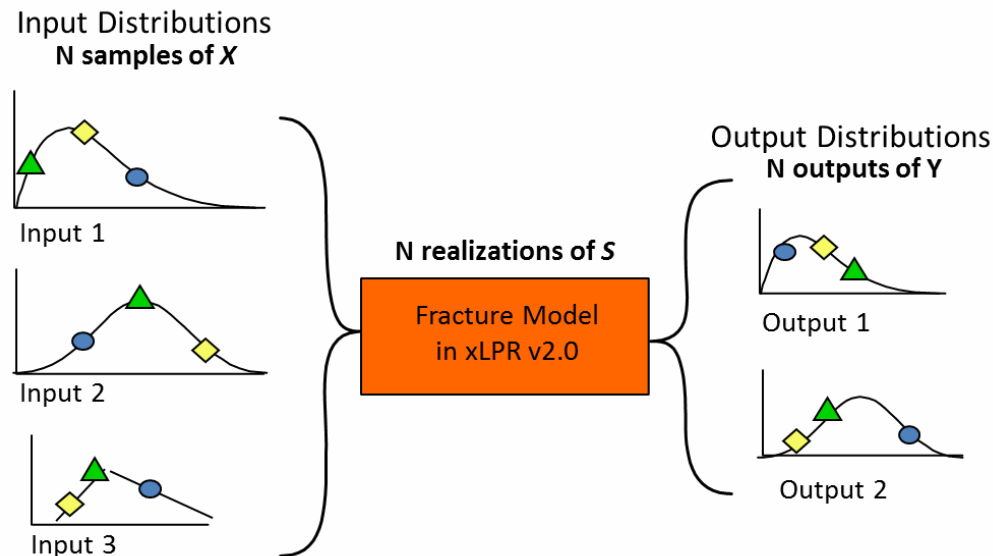


- **Probabilistic approach** evaluates the system at many input values and maps these to many outputs



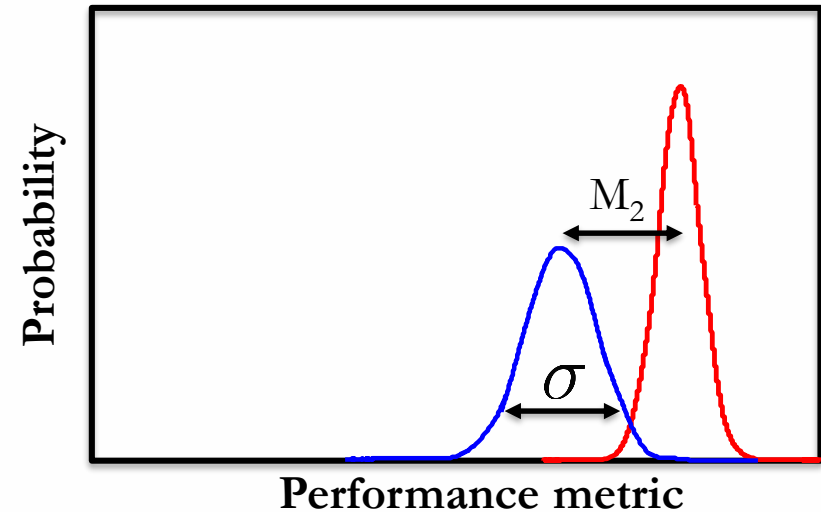
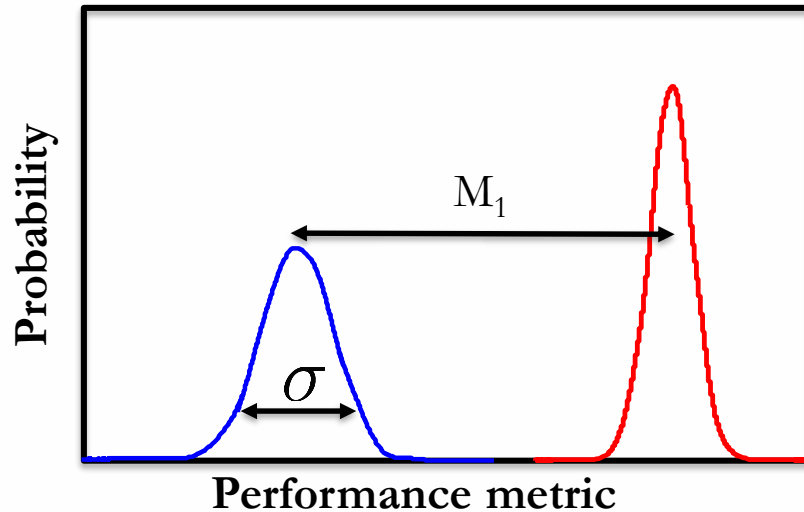


- **Steps for uncertainty propagation in PFM in xLPR v2.0:**
 1. Characterization of distributions on the uncertain input values
 2. Generation of samples from those distributions
 3. Propagation of samples through repeated fracture model execution
 4. Generate output distributions by repeating steps 2 and 3 N times
 5. Presentation of uncertainty analysis results in the form of functions of the output





MARGIN OF SAFETY: PROBABILISTIC APPROACH

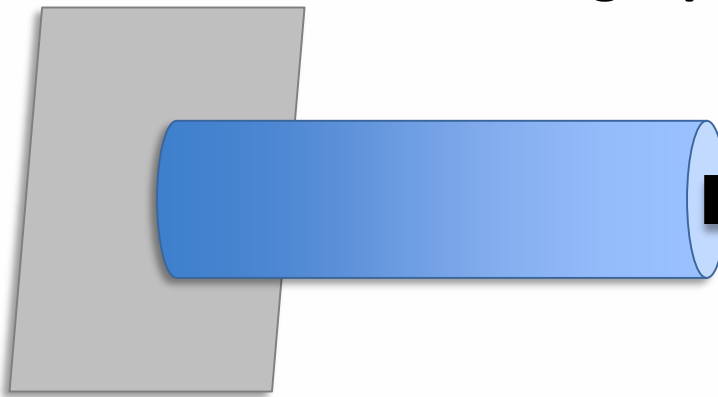


- **Probability distributions** are assigned to variables which have a significant effect on the problem (random variables)
- Problem is solved to determine **probability of desired results**



EXERCISE: PROBABILISTIC AXIAL STRESS IN A BAR

Example Problem: Calculate the stress in an axially loaded bar and determine if it will fail using a probabilistic approach



$$T \sim N(15000, 5000) \text{ [Ksi]}$$

$$r \sim N(0.8, 0.2) \text{ [in]}$$

Normal
Distribution

Mean

Standard
Deviation

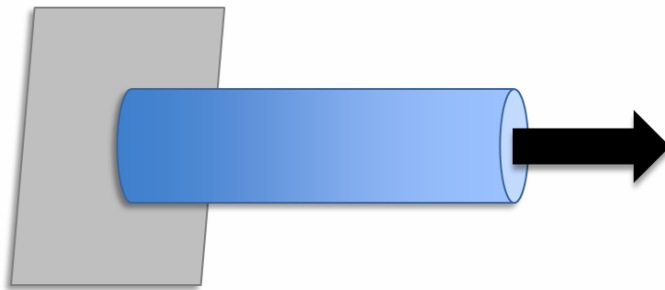
- Flow stress: $\sigma_f \sim N(55, 15) \text{ [Ksi]}$

Notation: This mathematical expression is read as “The flow stress is approximately Normally distributed with a mean of 55 Ksi and a standard deviation of 15 Ksi.”



EXERCISE: PROBABILISTIC AXIAL STRESS IN A BAR

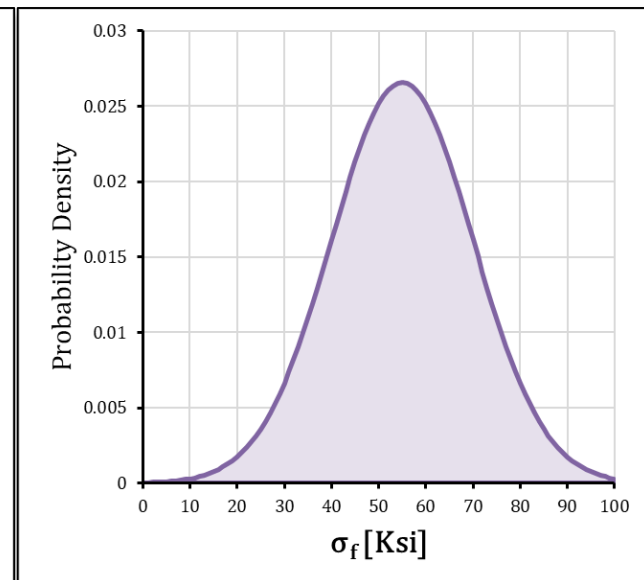
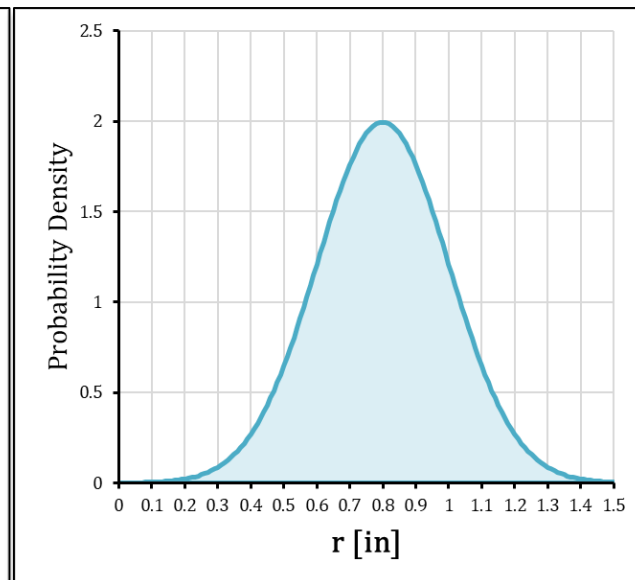
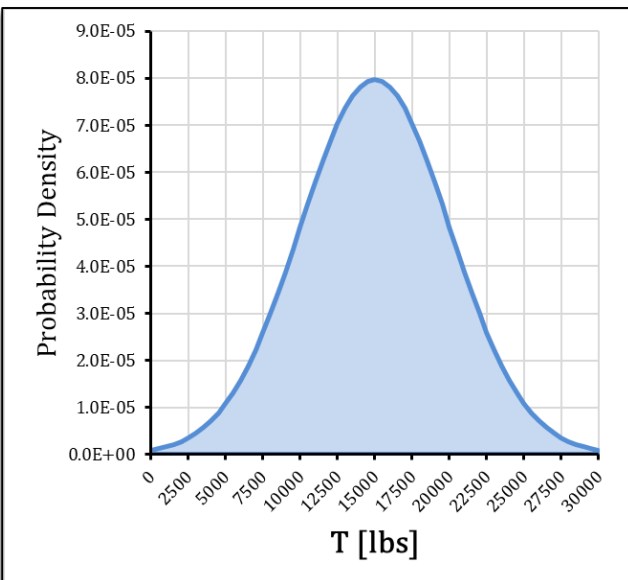
Example Problem: Calculate the stress in an axially loaded bar and determine if it will fail using a probabilistic approach



$$T \sim N(15000, 5000) \text{ [Ksi]}$$

$$r \sim N(0.8, 0.2) \text{ [in]}$$

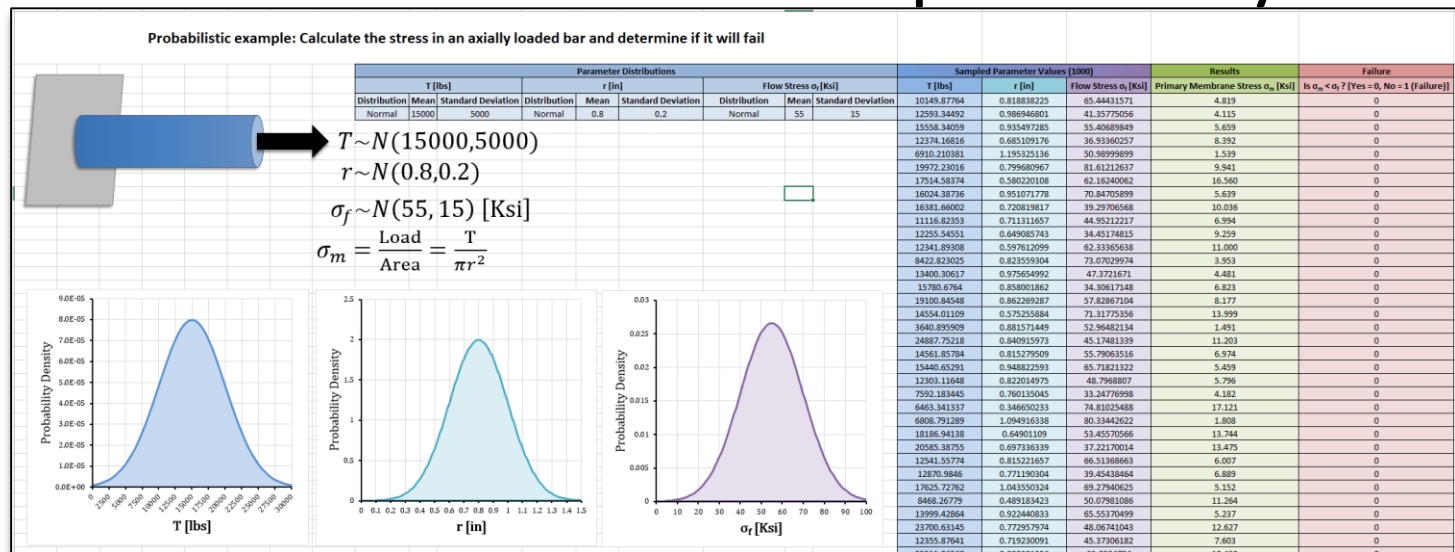
$$\sigma_f \sim N(55, 15) \text{ [Ksi]}$$





EXERCISE: PROBABILISTIC AXIAL STRESS IN A BAR

- Navigate to \Exercises\Module 1 and open the file “Module_1_Axial_Stress_in_a_Bar.xlsx”
- On the second tab, “Probabilistic approach”, enter the distribution values
- Does the bar fail? What is the probability of failure?

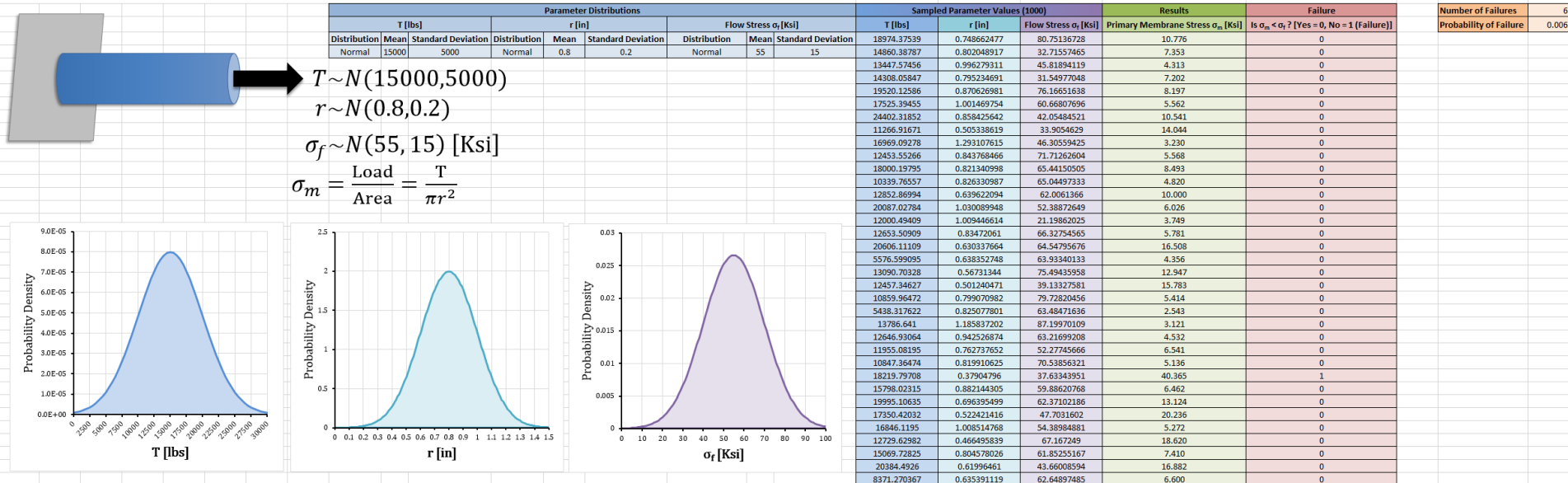




EXERCISE: PROBABILISTIC AXIAL STRESS IN A BAR

- **Solution:** Does the bar fail?
 - The bar experiences 6 failures in the screenshot below.
 - The calculated probability of failure is $6/1000 = 0.006$
 - Press 'F9' to rerun the calculation and get a different result

Probabilistic example: Calculate the stress in an axially loaded bar and determine if it will fail





EXERCISE: PROBABILISTIC AXIAL STRESS IN A BAR

- **Challenge problem:** Using the Excel file, fill in the table below by adding samples to the calculation and changing the mean of T
 - Create copies of the probabilistic tab for each row

Number of Samples	Mean of T [lbs]	Number of Failures	Probability of Failure
1000	15,000		
5000	15,000		
10,000	15,000		
10,000	20,000		
10,000	25,000		

- Which of these failure probabilities changes the most when you recalculate with new samples (Press F9)?



EXERCISE: PROBABILISTIC AXIAL STRESS IN A BAR

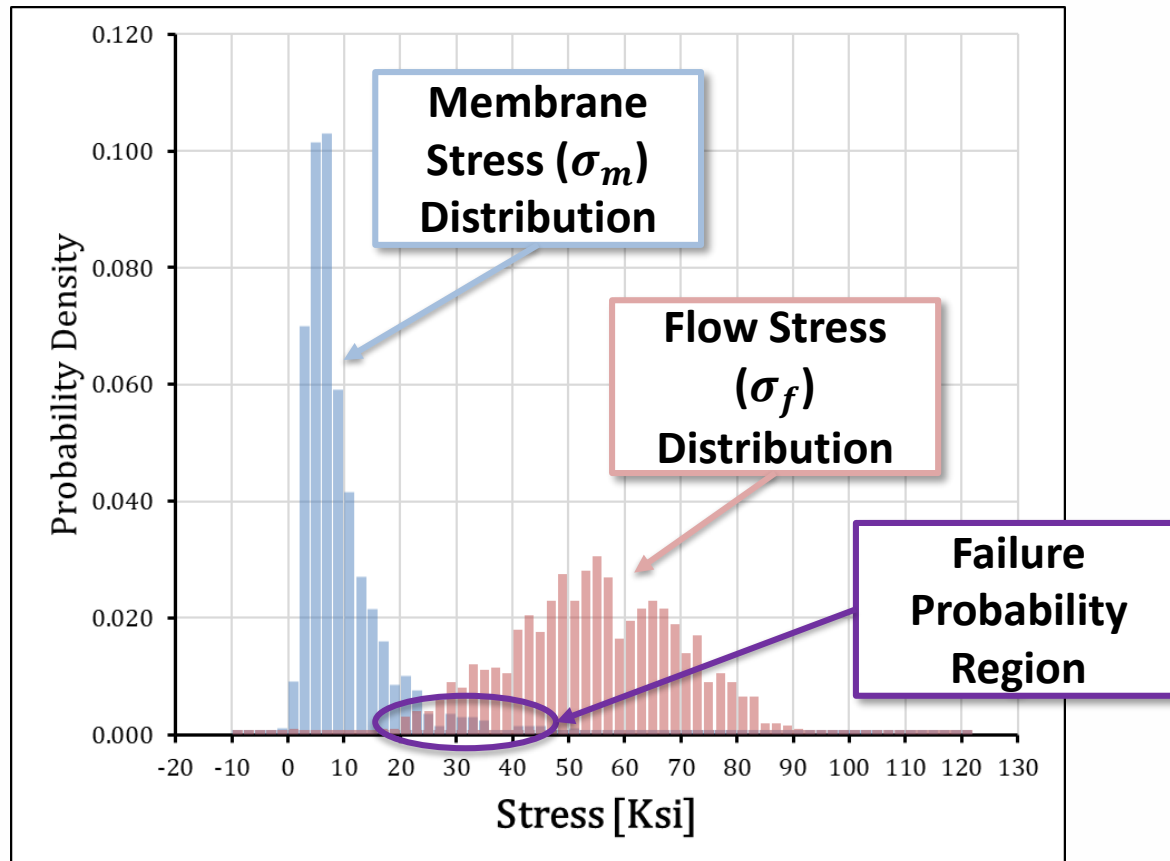
- **Challenge problem solution:**

Number of Samples	Mean of T [lbs]	Number of Failures	Probability of Failure
1000	15,000	~8	~0.08
5000	15,000	~59	~0.0118
10,000	15,000	~107	~0.0107
10,000	20,000	~197	~0.0197
10,000	25,000	~315	~0.315

- With fewer samples, the estimate of the probability of failure is less precise.

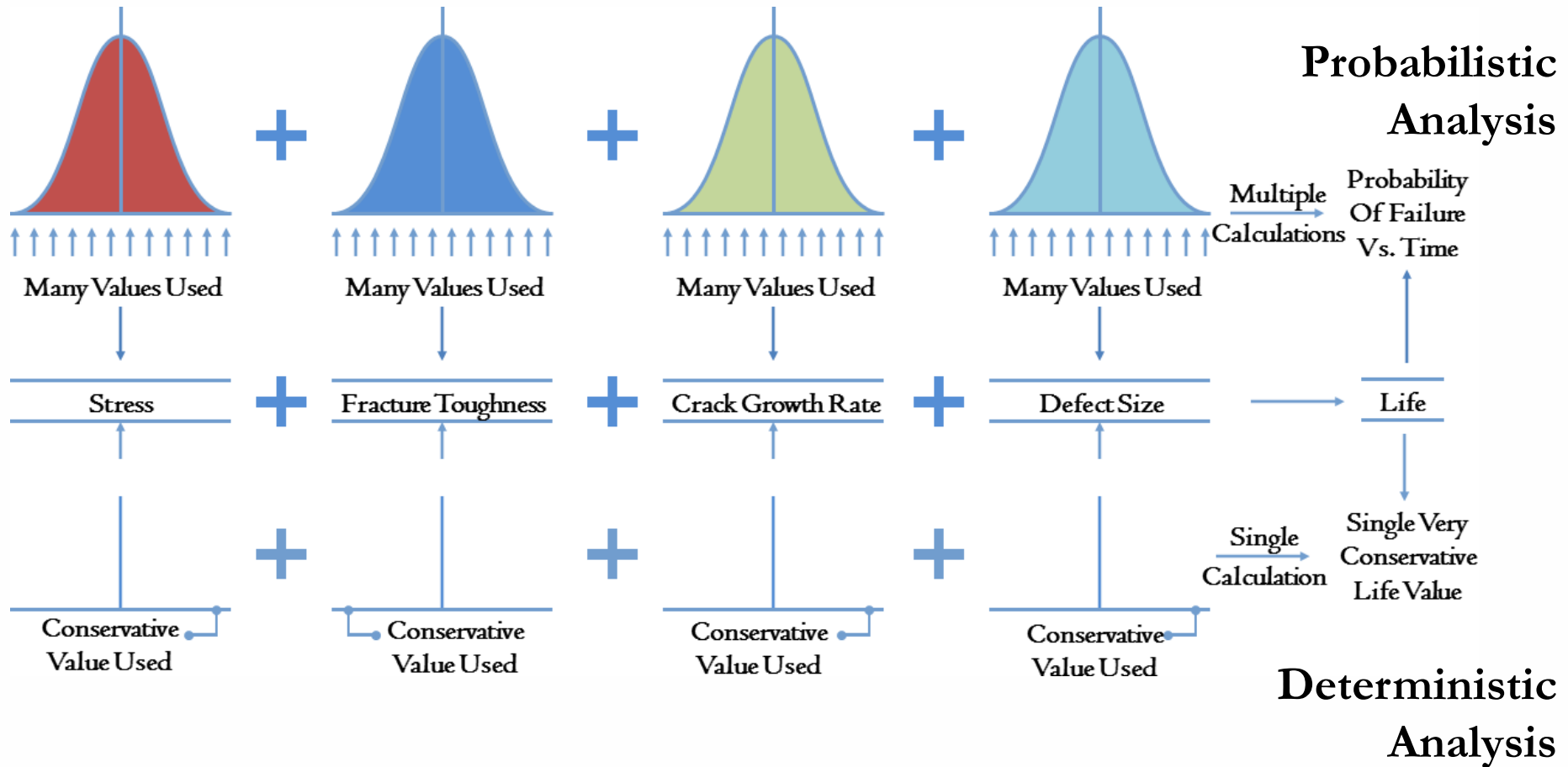


- Probabilistic approach results in probability of failure region





COMPARISON OF DETERMINISTIC AND PROBABILISTIC APPROACHES





Components of a PFM Analysis

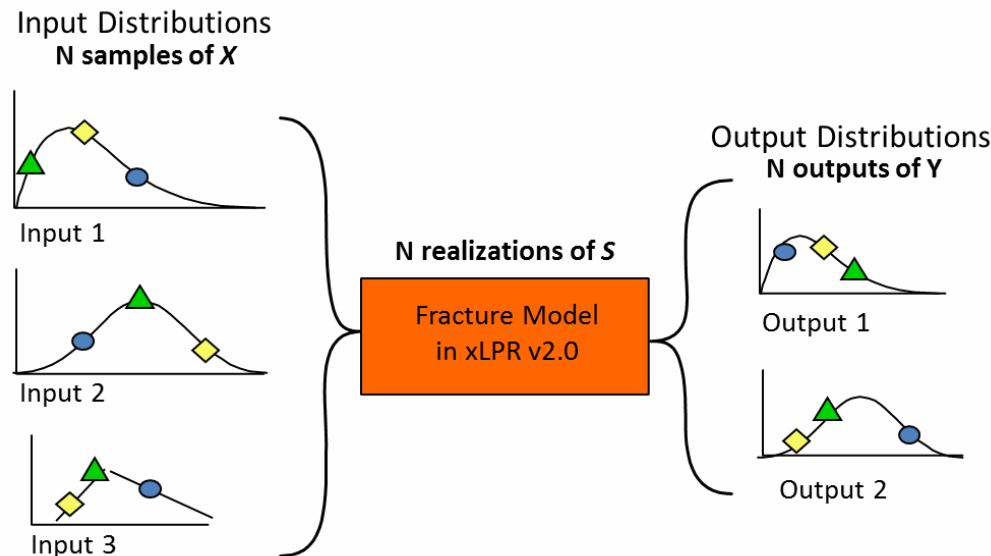


SOURCES OF UNCERTAINTY IN DFM AND PFM

- Uncertainties exist in both **Deterministic Fracture Mechanics (DFM)** models and **Probabilistic Fracture Mechanics (PFM)** models:
 - **Model uncertainty:** The **fracture model structure**, i.e., how accurately the deterministic fracture model describes the actual fracture process
 - **Numerical scheme uncertainty :** The **numerical approximation**, i.e., how appropriately the numerical method is used to approximating the model
 - **Input/Output uncertainty:** May only be known approximately, may vary between instances of the mechanism/sub-model for which predictions are sought
- PFM also involves sampling uncertainty:
 - **Sampling uncertainty:** quantities of interest are uncertain because the model can only be run a finite number of times



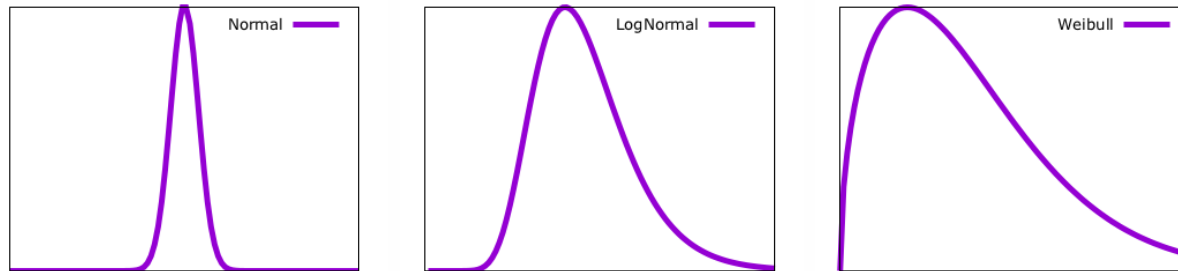
- **Steps for uncertainty propagation in PFM in xLPR v2.0:**
 1. **Characterization of distributions on the uncertain input values**
 2. Generation of samples from those distributions
 3. Propagation of samples through repeated fracture model execution
 4. Generate output distributions by repeating steps 2 and 3 N times
 5. Presentation of uncertainty analysis results in the form of functions of the output





- Input uncertainty is usually characterized using **probability distribution functions**

- Distribution will depend on the context of the application (e.g., one weld vs. a collection of welds)



- Some combinations of inputs are statistically dependent, meaning that sampling them independently could lead to non-physical conditions
 - Example: Inner diameter of a pipe must be greater than outer diameter
 - Correlations can be used to characterize the dependency between inputs to avoid



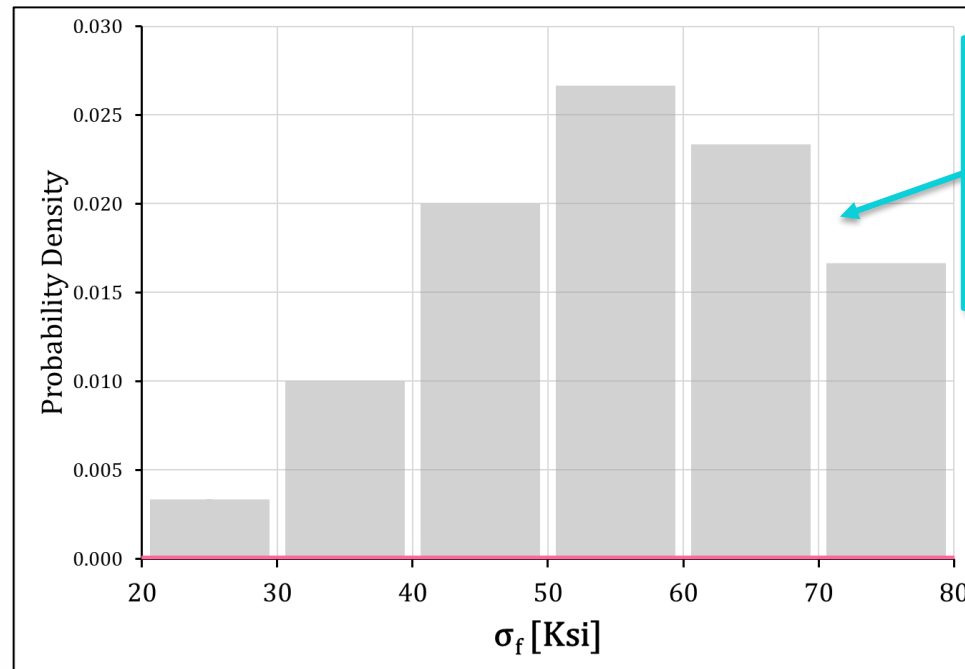
- Traditional techniques to **select probability distributions** include:
 - **Expert review:** used when no data is available
 - **Bayesian update:** used when data becomes available
 - **Distribution fitting:** enough data available to fit distribution
- Other techniques include: evidence theory, special objective response surface, etc.
- In the context of xLPR v2.0, much of the input calibration was done through distribution fitting using field and lab data

↓
Data availability
increases



EXERCISE: DISTRIBUTION FITTING

- Navigate to \Exercises\Module 1 and open the file “Module_1_Distribution_Fitting.xlsx”
- On the first tab, “Distribution Fitting”, we will fit distributions using 30 measurements of flow stress



**Empirical
distribution derived
from 30
measurements**



EXERCISE: DISTRIBUTION FITTING

- We will fit **Lognormal** and **Normal** distributions to this data
- Lognormal Distribution
 - Could be parameterized in several ways depending on the software
 - In Excel, the log mean and log standard deviation are used
 - In GoldSim, either the true mean and true standard deviation or the geometric mean and geometric standard deviation are used
 - Enter the following equations to calculate these parameters:

Lognormal Parameters	
mean($\ln(\sigma_i)$) [Log Mean]	=AVERAGE(T4:T33)
sd($\ln(\sigma_i)$) [Log SD]	=STDEV(T4:T33)
e ^{mean($\ln(\sigma_i)$)} [Geometric Mean]	=EXP(AVERAGE(T4:T33))
e ^{sd($\ln(\sigma_i)$)} [Geometric SD]	=EXP(STDEV(T4:T33))

Used in Excel

Used in GoldSim
(included here for information only)



EXERCISE: DISTRIBUTION FITTING

- Normal Distribution
 - Parameterized using the mean and standard deviation
 - Enter the following equations to calculate these parameters:

Normal Parameters	
mean(σ_i) [True Mean]	=AVERAGE(S4:S33)
sd(σ_i) [True SD]	=STDEV(S4:S33)

True mean and true standard deviation could also be used in GoldSim for the Lognormal distribution

- Distribution fitting for other distribution forms (i.e., Weibull) could also be accomplished using Maximum Likelihood Estimation (MLE)

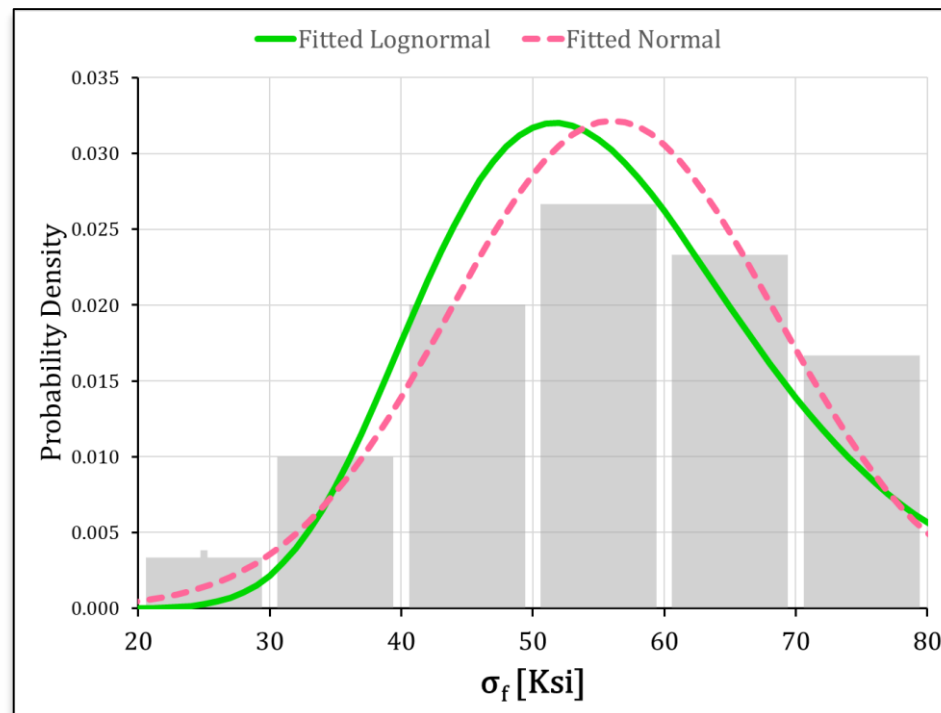


EXERCISE: DISTRIBUTION FITTING

- Final fitted distributions should look like this:

	Lognormal Parameters
$\text{mean}(\ln(\sigma_i))$ [Log Mean]	4.00
$\text{sd}(\ln(\sigma_i))$ [Log SD]	0.23
$e^{\text{mean}(\ln(\sigma_i))}$ [Geometric Mean]	54.63
$e^{\text{sd}(\ln(\sigma_i))}$ [Geometric Mean]	1.26

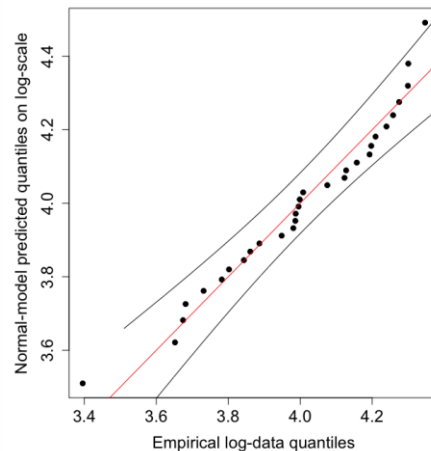
	Normal Parameters
$\text{mean}(\sigma_i)$ [True Mean]	56.04
$\text{sd}(\sigma_i)$ [True SD]	12.40



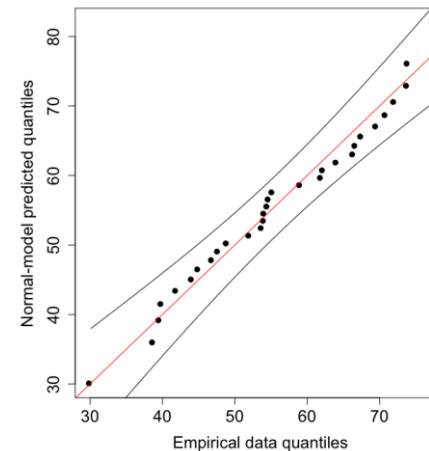


EXERCISE: DISTRIBUTION FITTING

- Statistical goodness of fit tests show that both of these distributions are consistent with the data
- QQ-plots with confidence intervals show the comparison between the data and fitted distributions
 - Red line represents a perfect fit, black dots are data points



Lognormal



Normal



EXERCISE: DISTRIBUTION FITTING

- Go to the “Probabilistic approach (1000)” tab
- Toggle between the Lognormal and Normal fits for flow stress to see how this impacts your results

Flow Stress σ_f [Ksi]		
Distribution	Mean	Standard Deviation
Lognormal	0.001	0.234448917
Normal		
Lognormal		

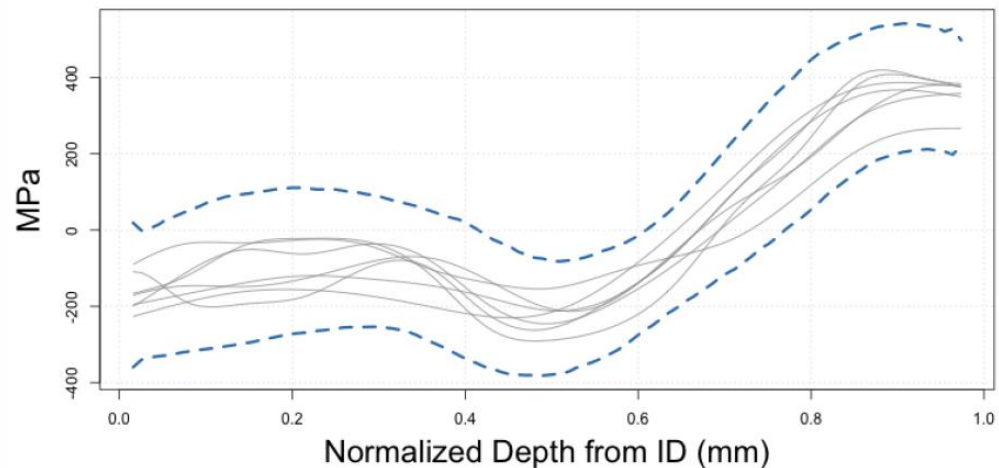
- Takeaways:
 - If we applied a more rigorous uncertainty analysis to this example we would show that failure probability calculations are sensitive to distribution tails
 - Sensitivity studies examining distribution choice should be applied in cases where the correct distribution is not clear



QUANTIFYING INPUT UNCERTAINTY: HOW TO GENERATE THESE DISTRIBUTIONS?

- Through **data from fracture experiments** on multiple pieces of hardware
 - **Pros:** Best way to quantify unit-to-unit variability
 - **Cons:** Expensive
- Through **field data** from legacy fractured systems
 - **Pros:** Real applicability
 - **Cons:** Knowledge from past might not be relevant to the future
- Through the **use of fracture models** representing the system
 - **Pros:** In principle, model can be executed many times to quantify uncertainty
 - **Cons:** Not always an accurate representation of reality
- A combination of both experiments and models, combining the pros and cons of each

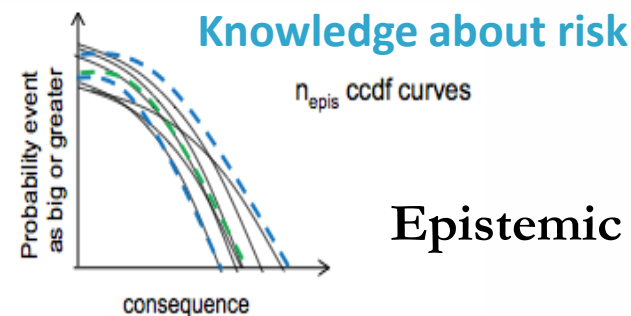
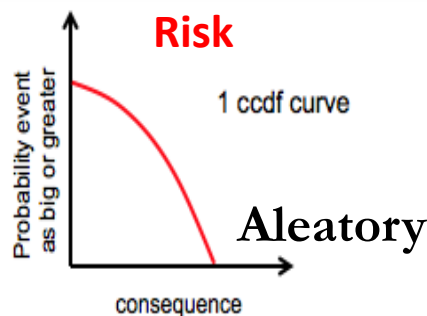
Example: field data and expert opinion are used to bound the model input weld residual stress





CLASSIFICATION OF UNCERTAINTY: ALEATORY VS. EPISTEMIC

- Model predictions are affected by two primary sources of uncertainty that can be used to classify inputs
 - Aleatory Uncertainty:** Natural (intrinsic) variability in system inputs and properties, considered irreducible
 - Example: Manufacturing differences among units
 - Epistemic Uncertainty:** Lack of knowledge about the system of interest, could be reduced by collecting additional information
 - Example: Uncertainty in crack growth model due to lack of data





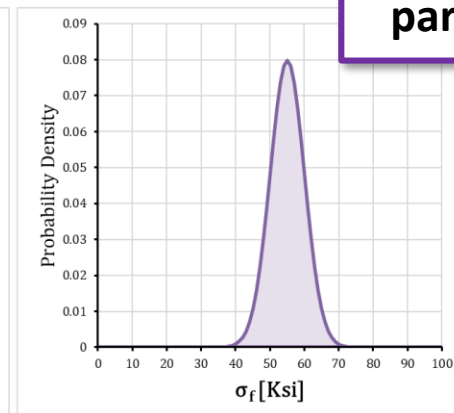
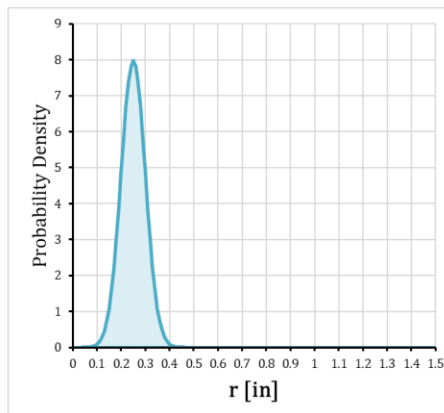
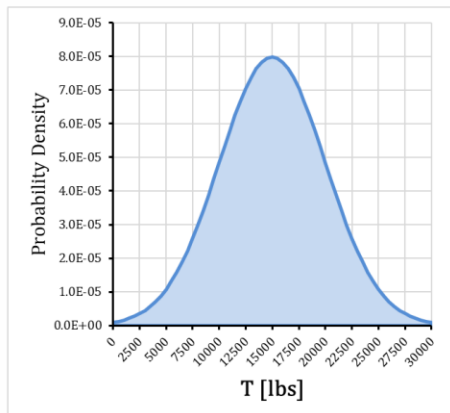
DEMONSTRATION: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- Navigate to \Exercises\Module 1 and open the file “Module_1_Separation_of_Uncertainty.xlsx”
- Open the second tab, “Basic Separation of Uncertainty”



$$\begin{aligned}T &\sim N(15000, 5000) \text{ (Epistemic)} \\r &\sim N(0.25, 0.05) \text{ (Epistemic)} \\ \sigma_f &\sim N(55, 5) \text{ [Ksi] (Aleatory)} \\ \sigma_m &= \frac{\text{Load}}{\text{Area}} = \frac{T}{\pi r^2}\end{aligned}$$

Same bar problem with slightly different distribution parameters





DEMONSTRATION: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

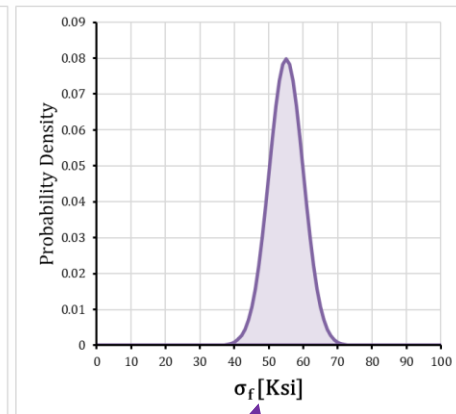
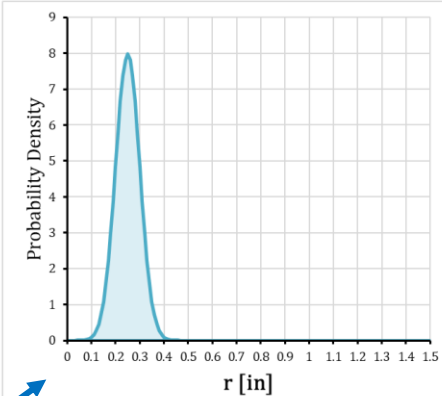
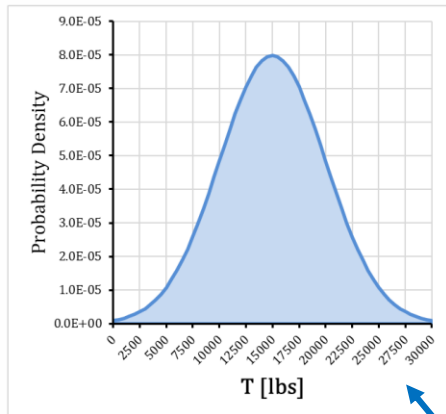


$$T \sim N(15000, 5000) \text{ (Epistemic)}$$

$$r \sim N(0.25, 0.05) \text{ (Epistemic)}$$

$$\sigma_f \sim N(55, 5) \text{ [Ksi] (Aleatory)}$$

$$\sigma_m = \frac{\text{Load}}{\text{Area}} = \frac{T}{\pi r^2}$$



T and r are set to epistemic:

Uncertainty in these inputs could be reduced by more measurements or more accurate measurement techniques

σ_f is set to aleatory:

Uncertainty in this input is due to random variability in material properties



DEMONSTRATION: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- What do these types of uncertainty mean in practice?
 - In this demonstration we have **3 epistemic** samples and **5 aleatory** samples and use simple random sampling
 - Example results shown below, press 'F9' in your Excel file to rerun the calculation

Aleatory samples

Epistemic samples

	Epistemic 1			Results	Failure	Epistemic 2			Results	Failure	Epistemic 3			Results	Failure
	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?
Aleatory 1	4674.356	0.263889	55.90232	21.366	0	12336.79	0.183541	53.0751509	116.570	1	13611.61204	0.284010218	56.88928	53.715	0
Aleatory 2	4674.356	0.263889	58.6355	21.366	0	12336.79	0.183541	55.59391582	116.570	1	13611.61204	0.284010218	48.34774	53.715	1
Aleatory 3	4674.356	0.263889	62.89263	21.366	0	12336.79	0.183541	57.79225462	116.570	1	13611.61204	0.284010218	60.04014	53.715	0
Aleatory 4	4674.356	0.263889	59.44506	21.366	0	12336.79	0.183541	48.4849068	116.570	1	13611.61204	0.284010218	54.65302	53.715	0
Aleatory 5	4674.356	0.263889	59.54206	21.366	0	12336.79	0.183541	45.02974798	116.570	1	13611.61204	0.284010218	55.84193	53.715	0

- What is the probability of failure?



DEMONSTRATION: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- The collection of **aleatory** samples for each epistemic sample represents an **estimate of the probability of failure**

	Epistemic 1			Results	Failure	Epistemic 2			Results	Failure	Epistemic 3			Results	Failure
	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?
Aleatory 1	4674.356	0.263889	55.90232	21.366	0	12336.79	0.183541	53.0751509	116.570	1	13611.61204	0.284010218	56.88928	53.715	0
Aleatory 2	4674.356	0.263889	58.6355	21.366	0	12336.79	0.183541	55.59391582	116.570	1	13611.61204	0.284010218	48.34774	53.715	1
Aleatory 3	4674.356	0.263889	62.89263	21.366	0	12336.79	0.183541	57.79225462	116.570	1	13611.61204	0.284010218	60.04014	53.715	0
Aleatory 4	4674.356	0.263889	59.44506	21.366	0	12336.79	0.183541	48.4849068	116.570	1	13611.61204	0.284010218	54.65302	53.715	0
Aleatory 5	4674.356	0.263889	59.54206	21.366	0	12336.79	0.183541	45.62974798	116.570	1	13611.61204	0.284010218	55.84193	53.715	0

For the first epistemic sample, the probability of failure is

$$\frac{0+0+0+0+0}{5} = 0$$

For the second epistemic sample, the probability of failure is

$$\frac{1+1+1+1+1}{5} = 1$$

For the first epistemic sample, the probability of failure is

$$\frac{0+1+0+0+0}{5} = 0.2$$



DEMONSTRATION: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- The collection of **epistemic** estimates of the probability of failure represents an estimate of the **distribution of the probability of failure**

	Epistemic 1			Results	Failure	Epistemic 2			Results	Failure	Epistemic 3			Results	Failure
	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?
Aleatory 1	4674.356	0.263889	55.90232	21.366	0	12336.79	0.183541	53.0751509	116.570	1	13611.61204	0.284010218	56.88928	53.715	0
Aleatory 2	4674.356	0.263889	58.6355	21.366	0	12336.79	0.183541	55.59391582	116.570	1	13611.61204	0.284010218	48.34774	53.715	1
Aleatory 3	4674.356	0.263889	62.89263	21.366	0	12336.79	0.183541	57.79225462	116.570	1	13611.61204	0.284010218	60.04014	53.715	0
Aleatory 4	4674.356	0.263889	59.44506	21.366	0	12336.79	0.183541	48.4849068	116.570	1	13611.61204	0.284010218	54.65302	53.715	0
Aleatory 5	4674.356	0.263889	59.54206	21.366	0	12336.79	0.183541	45.02974798	116.570	1	13611.61204	0.284010218	55.84193	53.715	0

Probability of failure = 0

Probability of failure = 1

Probability of failure = 0.2

- Is this a good estimate of the uncertainty in the probability of failure?



DEMONSTRATION: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- This sample size does not generate a stable estimate of the distribution of the probability of failure
 - Running the simulation again (push 'F9') creates very different results

	Epistemic 1			Results	Failure	Epistemic 2			Results	Failure	Epistemic 3			Results	Failure
	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?
Aleatory 1	4674.356	0.263889	55.90232	21.366	0	12336.79	0.183541	53.0751509	116.570	1	13611.61204	0.284010218	56.88928	53.715	0
Aleatory 2	4674.356	0.263889	58.6355	21.366	0	12336.79	0.183541	55.59391582	116.570	1	13611.61204	0.284010218	48.34774	53.715	1
Aleatory 3	4674.356	0.263889	62.89263	21.366	0	12336.79	0.183541	57.79225462	116.570	1	13611.61204	0.284010218	60.04014	53.715	0
Aleatory 4	4674.356	0.263889	65.12345	21.366	0	12336.79	0.183541	60.07567	116.570	1	13611.61204	0.284010218	62.35678	53.715	0
Aleatory 5	4674.356	0.263889	67.35678	21.366	0	12336.79	0.183541	62.35678	116.570	1	13611.61204	0.284010218	64.67890	53.715	0

Probability of failure = 0

Probability of failure = 1

Probability of failure = 0.2

	Epistemic 1			Results	Failure	Epistemic 2			Results	Failure	Epistemic 3			Results	Failure
	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?
Aleatory 1	13418	0.249613	54.32995	68.549	1	16266.9	0.314188	51.86511668	52.454	1	18942.78021	0.213831191	52.95519	131.872	1
Aleatory 2	13418	0.249613	57.73549	68.549	1	16266.9	0.314188	61.4391029	52.454	0	18942.78021	0.213831191	55.51452	131.872	1
Aleatory 3	13418	0.249613	54.02082	68.549	1	16266.9	0.314188	51.18712643	52.454	1	18942.78021	0.213831191	48.84877	131.872	1
Aleatory 4	13418	0.249613	57.73549	68.549	1	16266.9	0.314188	61.4391029	52.454	0	18942.78021	0.213831191	55.51452	131.872	1
Aleatory 5	13418	0.249613	54.02082	68.549	1	16266.9	0.314188	51.18712643	52.454	1	18942.78021	0.213831191	48.84877	131.872	1

Probability of failure = 1

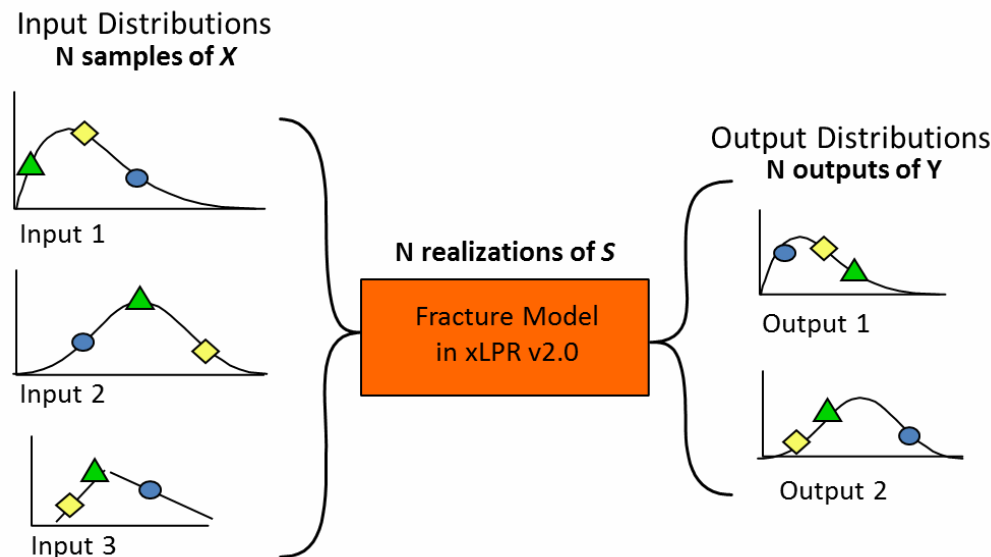
Probability of failure = 0.6

Probability of failure = 1

- Additional samples and techniques for uncertainty characterization and propagation may be needed

- Steps for PFM in xLPR v2.0:

1. Characterization of distributions on the uncertain input values
- 2. Generation of samples from those distributions**
3. Propagation of samples through repeated fracture model execution
4. Generate output distributions by repeating steps 2 and 3 N times
5. Presentation of uncertainty analysis results in the form of functions of the output



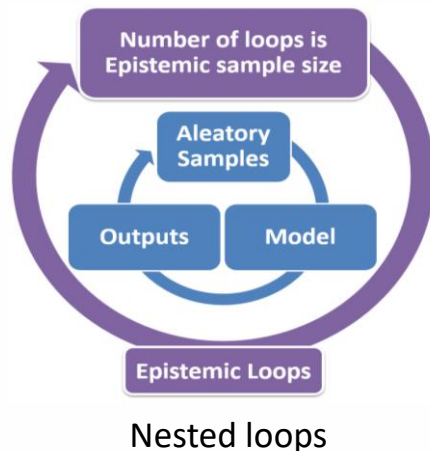


- When sampling inputs, there are **three decisions** to make:
 - Whether or not to **separate aleatory and epistemic** uncertainty
 - What **sampling scheme** to use for repeatedly sampling uncertain inputs
 - **How many samples** should be included in the sampling scheme

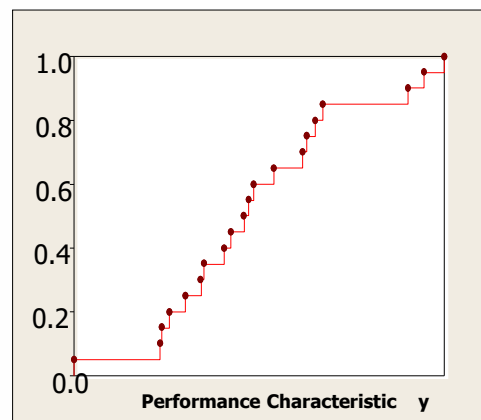


SEPARATION OF UNCERTAINTY

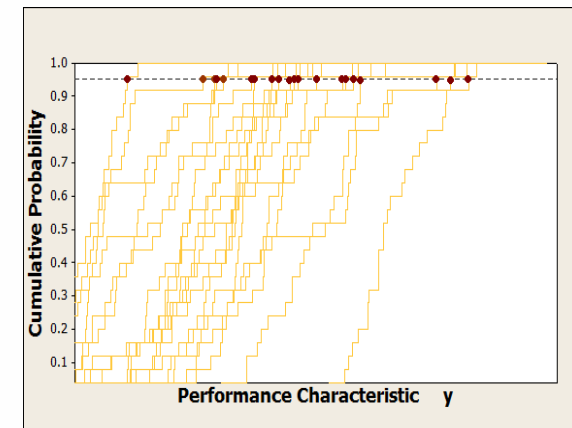
- Uncertainty in xLPR v2.0 can be **separated into aleatory and epistemic** using **nested sampling loops**
 - Requires greater computational power (for simulation and for quantifying uncertainty)
- **Separation** may facilitate **understanding of epistemic uncertainties** (knowledge of risk)
 - Refine distributions on the uncertain inputs that contribute most to the output uncertainty and repeat analysis.



Inner (aleatory) loop



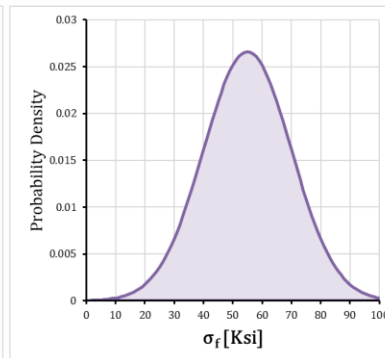
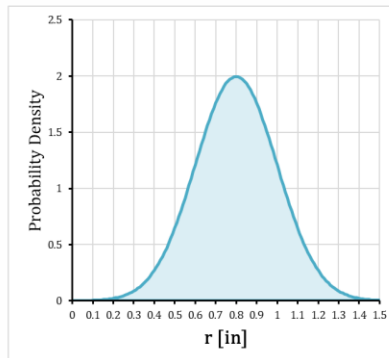
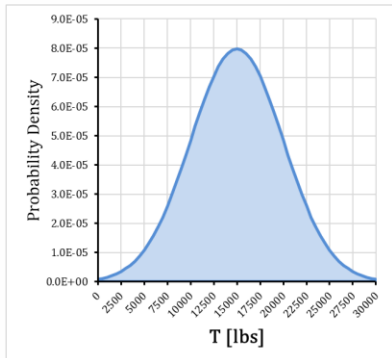
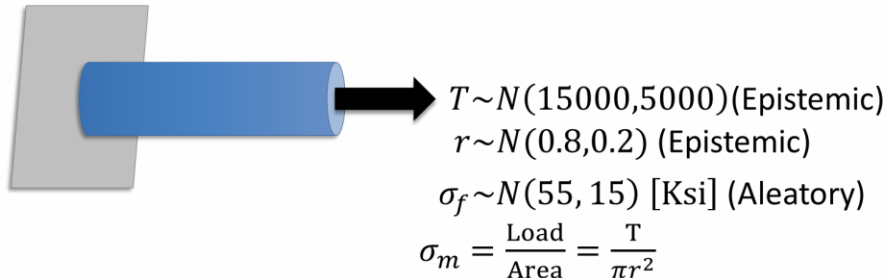
Outer (epistemic) loop





EXERCISE: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- Navigate to \Exercises\Module 1 and open the file “Module_1_Separation_of_Uncertainty.xlsx”
- Open the second tab, “Axial Stress in a Bar(100x50)”
 - Same bar problem with original distributions



T and r are set to epistemic:

Uncertainty in these inputs could be reduced by more measurements or more accurate measurement techniques

σ_f is set to aleatory:

Uncertainty in this input is due to random variability in material properties



EXERCISE: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- In this exercise we are using **100 epistemic** samples and **50 aleatory** samples with simple random sampling
- This expands on the number of samples demonstrated earlier in the presentation

Aleatory samples

Epistemic samples

	Epistemic 1			Results	Failure	Epistemic 2			Results	Failure	Epistemic 3			Results	Failure
	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?	T [lbs]	r [in]	σ_f [Ksi]	σ_m [Ksi]	Is $\sigma_m < \sigma_f$?
Aleatory 1	4674.356	0.263889	55.90232	21.366	0	12336.79	0.183541	53.0751509	116.570	1	13611.61204	0.284010218	56.88928	53.715	0
Aleatory 2	4674.356	0.263889	58.6355	21.366	0	12336.79	0.183541	55.59391582	116.570	1	13611.61204	0.284010218	48.34774	53.715	1
Aleatory 3	4674.356	0.263889	62.89263	21.366	0	12336.79	0.183541	57.79225462	116.570	1	13611.61204	0.284010218	60.04014	53.715	0
Aleatory 4	4674.356	0.263889	59.44506	21.366	0	12336.79	0.183541	48.4849068	116.570	1	13611.61204	0.284010218	54.65302	53.715	0
Aleatory 5	4674.356	0.263889	59.54206	21.366	0	12336.79	0.183541	45.02974798	116.570	1	13611.61204	0.284010218	55.84193	53.715	0

- Press 'F9' to run the calculation repeatedly and watch the estimates of the probability of failure change
- Is there a high level of variability in the estimate of the average probability of failure?



EXERCISE: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- Change the standard deviations distributions of T and r to match the table below:

Parameter Distributions								
T [lbs]			r [in]			Flow Stress σ_f [Ksi]		
Distribution	Mean	Standard Deviation	Distribution	Mean	Standard Deviation	Distribution	Mean	Standard Deviation
Normal	15000	100	Normal	0.8	0.01	Normal	55	15

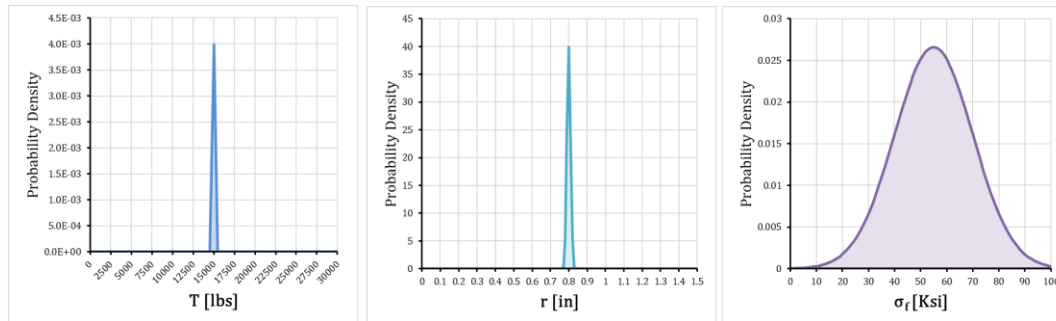
- Press 'F9' to run the calculation again and watch the estimates of the probability of failure change
 - Record each value by copying it as a value
 - into a new row
- Is there less variability in the estimate of average probability of failure?



EXERCISE: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- There is still a great deal of **aleatory sampling uncertainty**
- We can reduce this by adding aleatory samples
- Open the third tab, “Axial Stress in a Bar(100x500)”
 - The number of aleatory samples has been increased from 50 to 500

Epistemic distributions are still narrower



- Press ‘F9’ to run the calculation again and watch the estimates of the probability of failure change
- Is there less variability in the average probability of failure?

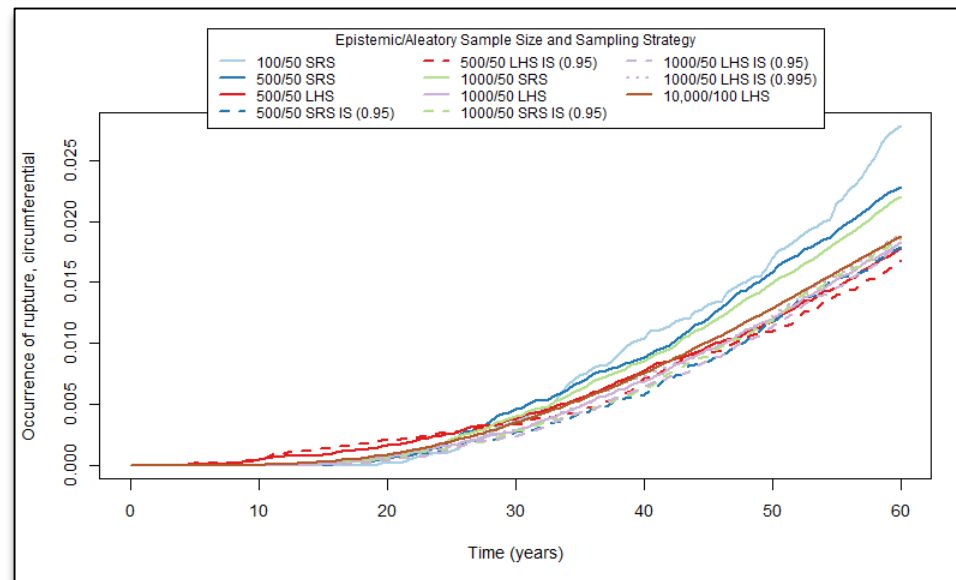


EXERCISE: SEPARATION OF EPISTEMIC AND ALEATORY UNCERTAINTY

- Exercise takeaways:
 - Variability in the estimate of the average probability of failure in this simple example was **reduced** by:
 - Reducing uncertainty in input distributions – this is not always possible in real application
 - Increasing the aleatory sample size
 - If uncertainty in the epistemic input distributions could not be reduced, we could have also increased the epistemic sample size
 - Advanced sampling techniques could also be applied to reduce this variability



- The number of Monte Carlo samples is finite, we cannot run an infinite number of realizations to achieve exact answer
 - Therefore, quantities of interest (QoIs)(i.e., mean probability of rupture) have sampling uncertainty
 - Example: a QoI estimate will be more accurate using 1,000,000 samples than 100 samples



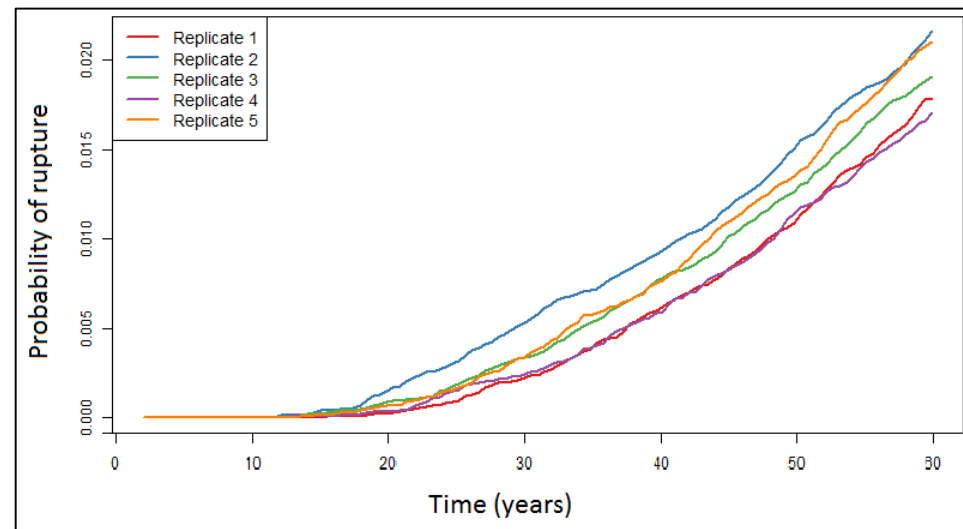
Example: Mean probability of occurrence of circumferential rupture for Scenario 3, Runs 1 through 10.



- Sampling schemes should be selected such that sampling uncertainty in the QoI is sufficiently small
 - Increases in sample size or in the sampling algorithm can be used to accomplish this

Running a model 5 different times results in 5 different QoI estimates

Differences in the estimates are due to sampling uncertainty





- There are many ways to sample inputs for uncertainty propagation:
 - Simple Random Sampling (SRS)
 - Latin Hypercube Sampling (LHS)
 - Importance Sampling (IS) [Can be used with LHS or SRS]
 - Discrete Probability Distribution (DPD)
 - Adaptive Sampling [Can be used with LHS or SRS]
 - Other methods (quasi-MC, etc.)
 - Alternative to sampling-based methods (FORM, SORM...)





- There are many ways to sample inputs for uncertainty propagation:

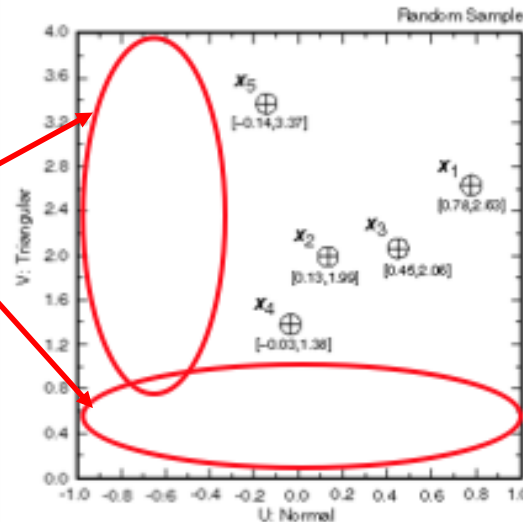
- Simple Random Sampling (SRS)
- Latin Hypercube Sampling (LHS)
- Importance Sampling (IS) [Can be used with LHS or SRS]
- Discrete Probability Distribution (DPD)
- Adaptive Sampling [Can be used with LHS or SRS]
- Other methods (quasi-MC, etc.)
- Alternative to sampling-based methods (FORM, SORM...)



**Available in
xLPR**

- The simplest Monte Carlo sampling scheme is simple random sampling (SRS)
 - All inputs are randomly sampled from their input distributions
 - Pros:** easy to implement, easy to explain, easy to analyze data
 - Cons:** sufficiently large samples may not be possible to achieve reasonably low sampling uncertainty

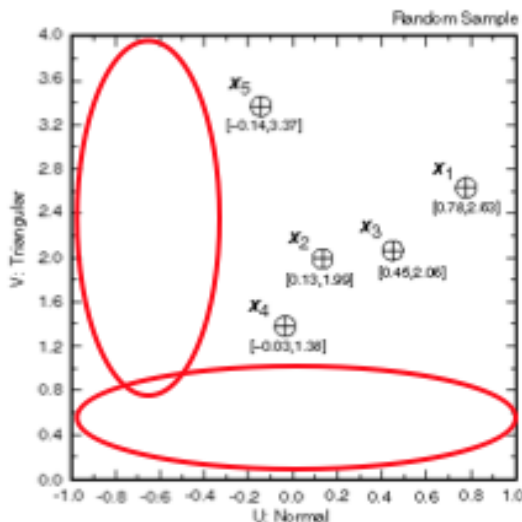
Parts of the input space might be missed



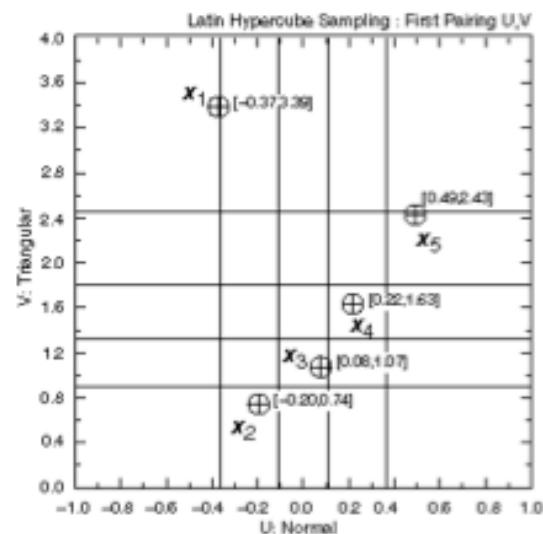
SRS



- LHS “forces” samples to be spread out across domain of the input distributions using dense stratification across the range of each variable
 - **Pros:** lower sampling uncertainty than SRS, easy to analyze
 - **Cons:** more difficult to estimate sampling uncertainty



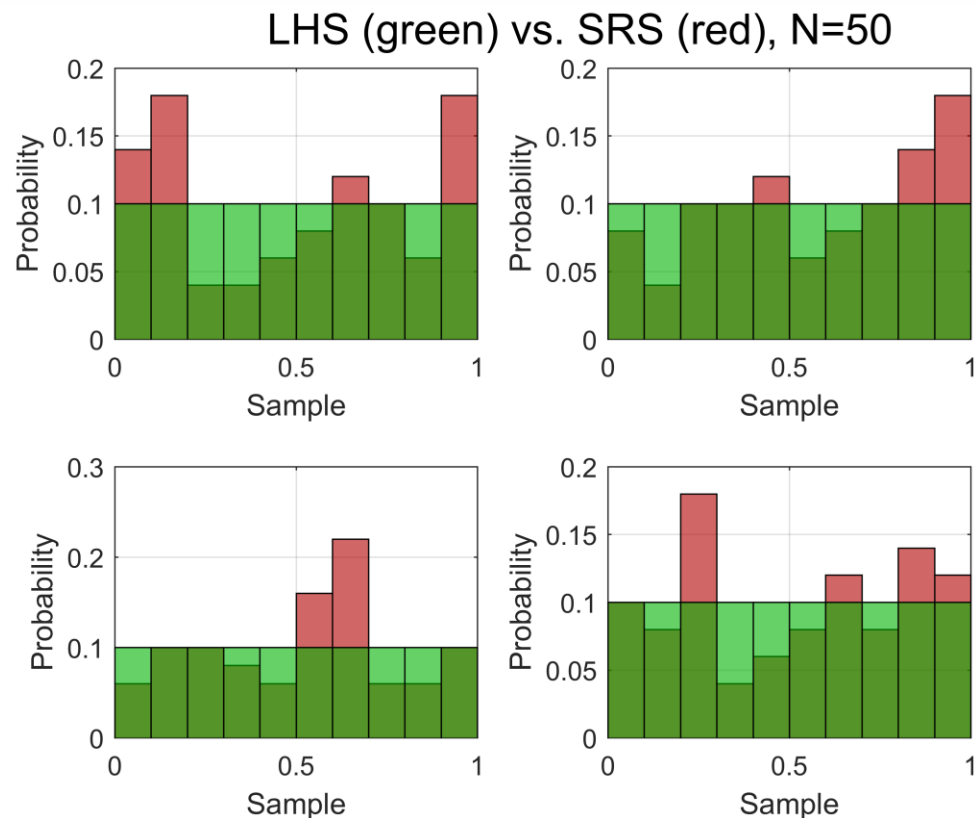
SRS



LHS



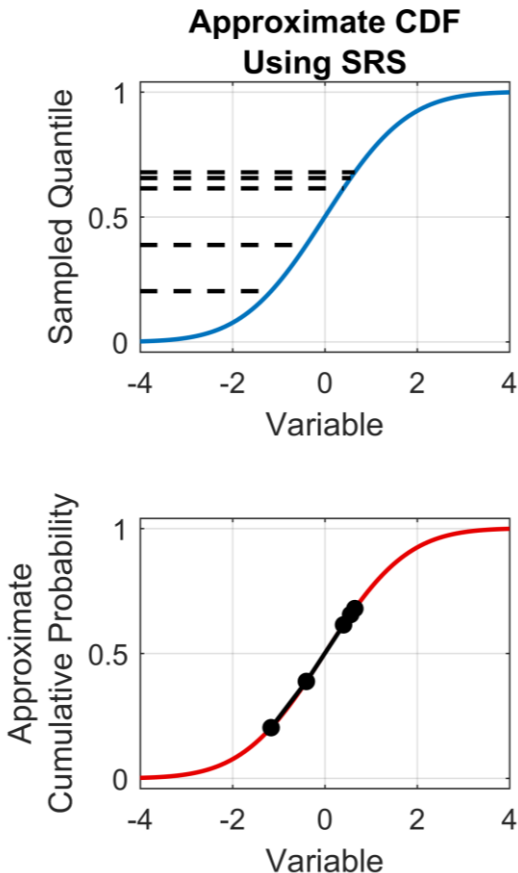
- LHS covers the range of all inputs while this coverage is less consistent under SRS





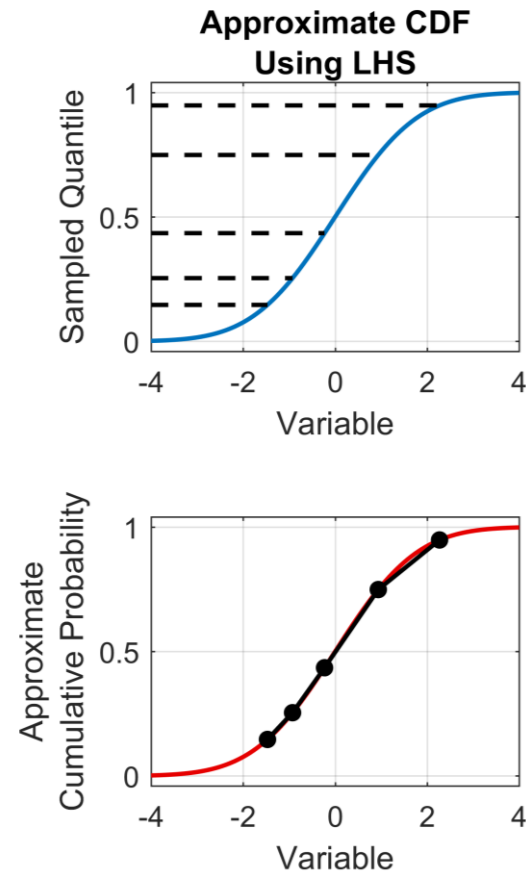
SRS VS LHS: ILLUSTRATION

SRS



(x_1, \dots, x_5) are the result of drawing a random sequence. Note the clustering and gaps that result from the sample.

LHS

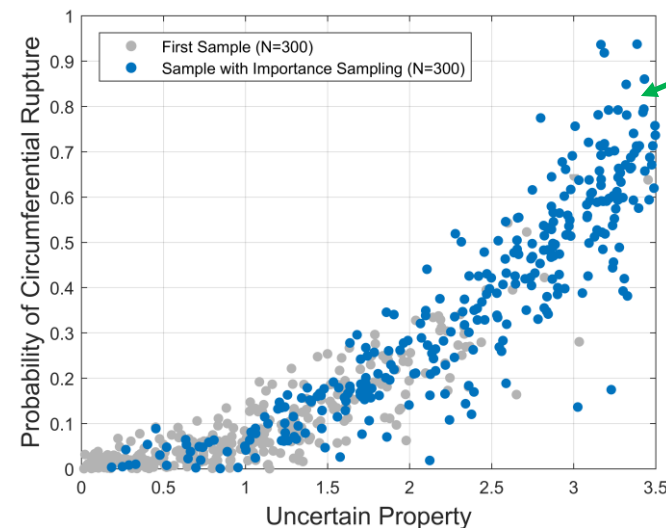
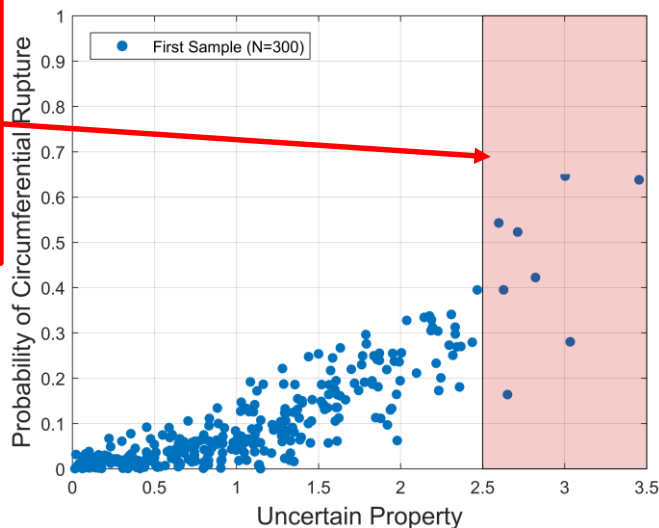


For n samples, we use the partition $[0, 1/n, 2/n, \dots, (n-1)/n, 1]$ on the cumulative probability axis with one sample in each partition. Note the even coverage of the CDF.



- Over-sample ‘important’ parts of the input space
 - **Pros:** better estimation of rare event probabilities
 - **Cons:** harder to implement, more difficult to analyze data, bad implementation can increase sampling uncertainty
- Results are calculated using **importance sampling weights** so that this sampling scheme doesn’t skew the output

Area of interest not covered by initial sampling



Importance sampling allows estimation of area of interest



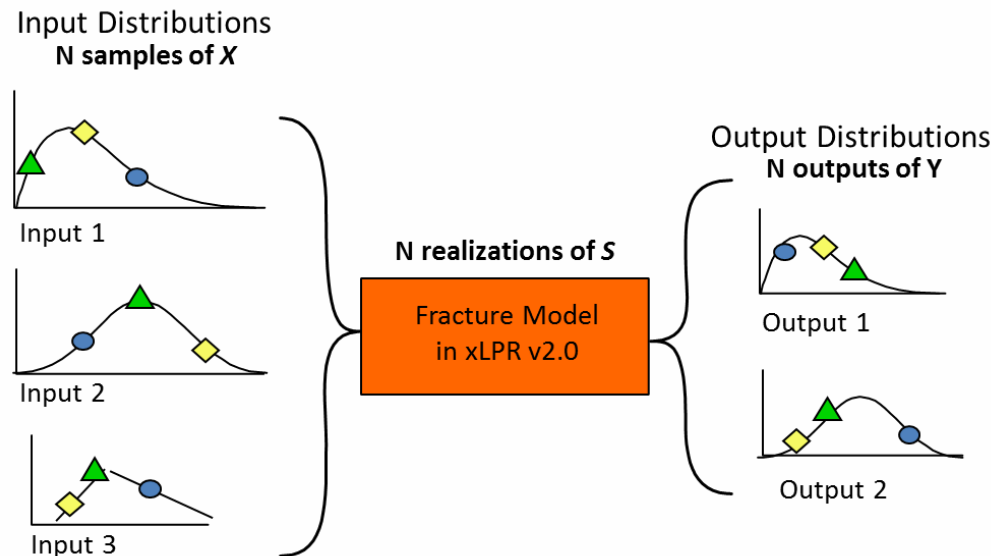
- **Simple random sampling** is simplest – easy to analyze, combine results across runs, calculate sampling uncertainty
- **Latin hypercube sampling** is an improvement on simple random sampling without increasing the computation time or complexity of post-processing
- **Importance sampling** helps estimate very small probabilities in reasonable computing time
 - This type of sampling scheme is chosen after preliminary sensitivity studies have been conducted



- **Steps for PFM in xLPR v2.0:**

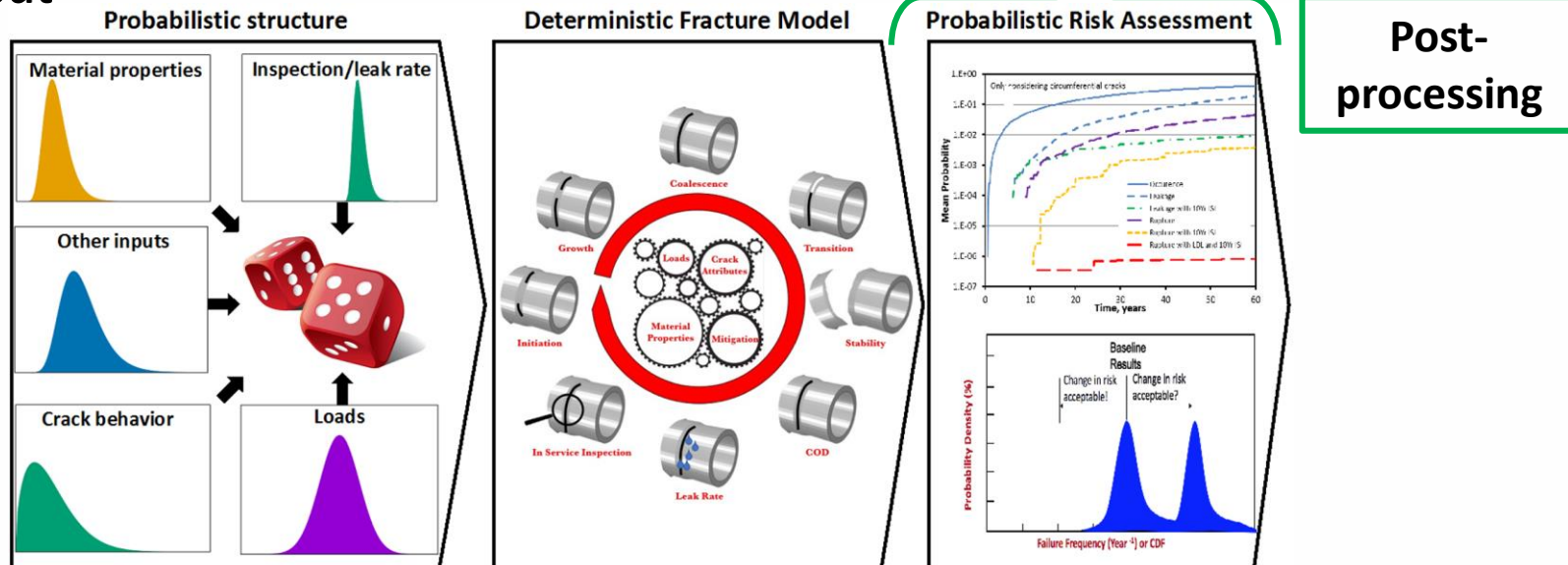
1. Characterization of distributions on the uncertain input values
2. Generation of samples from those distributions
- 3. Propagation of samples through repeated fracture model execution**
- 4. Generate output distributions by repeating steps 2 and 3 N times**
5. Presentation of uncertainty analysis results in the form of functions of the output

Accomplished
using the
xLPR
Framework



- Steps for PFM in xLPR v2.0:

1. Characterization of distributions on the uncertain input values
2. Generation of samples from those distributions
3. Propagation of samples through repeated fracture model execution
4. Generate output distributions by repeating steps 2 and 3 N times
5. **Presentation of uncertainty analysis results in the form of functions of the output**

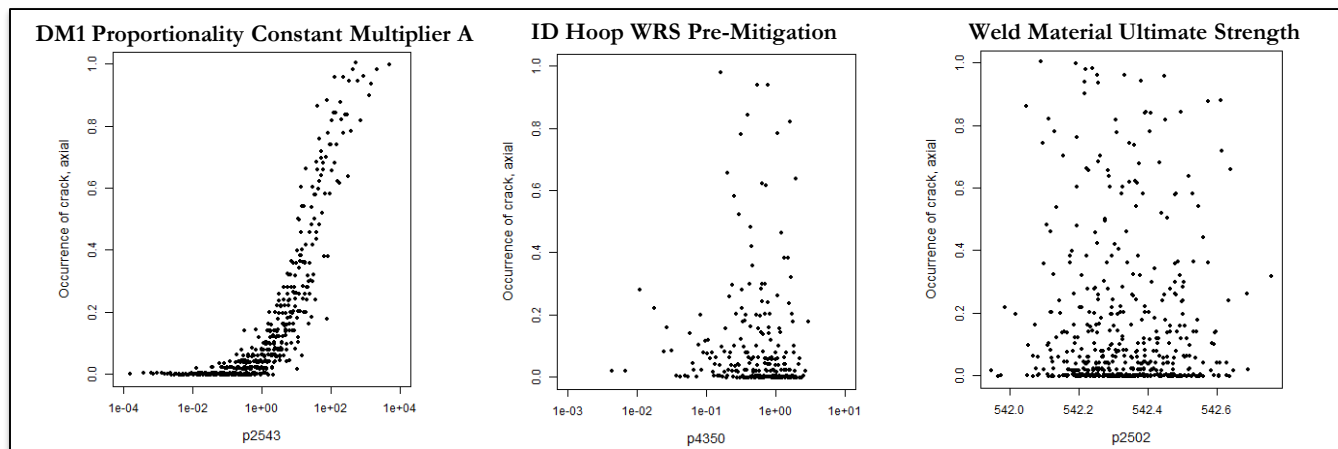




- **Sensitivity analysis:** Estimate the relationship of the uncertainty in the inputs to the uncertainty in the outputs of interest
 - Identify the inputs that contribute to the majority of the uncertainty in the outputs

Probability of Occurrence of Axial Cracks – $R^2 = 0.892$		
Variable Identifier	Variable Name	SRRC
p2543	Multiplier proport. Const. A (DM1)	0.902
p4350	Hoop WRS Pre-mitigation	-0.228
p5103	Log reg. intercept param., beta_0 (axial)	0.064
p5104	Log reg. slope param., beta_1 (axial)	0.050
p1102	Pipe wall thickness	-0.039

Example: Five most important variables identified using linear rank regression.

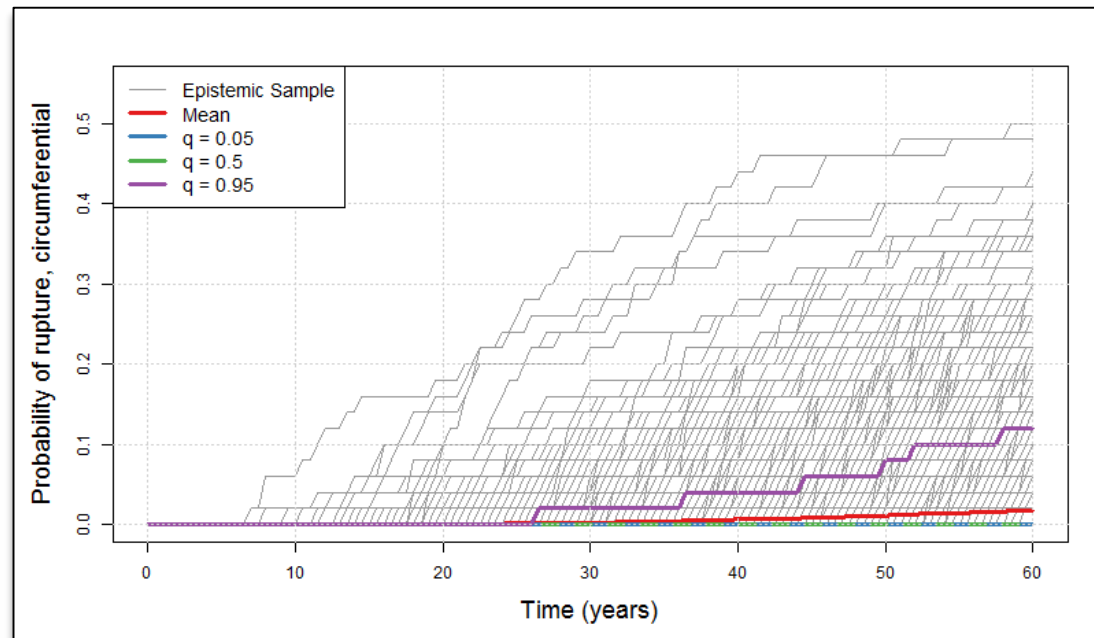


Example: Scatter plots showing relationship between sampled variables and output of interest.



- **Uncertainty analysis:** Analysis of the output uncertainty

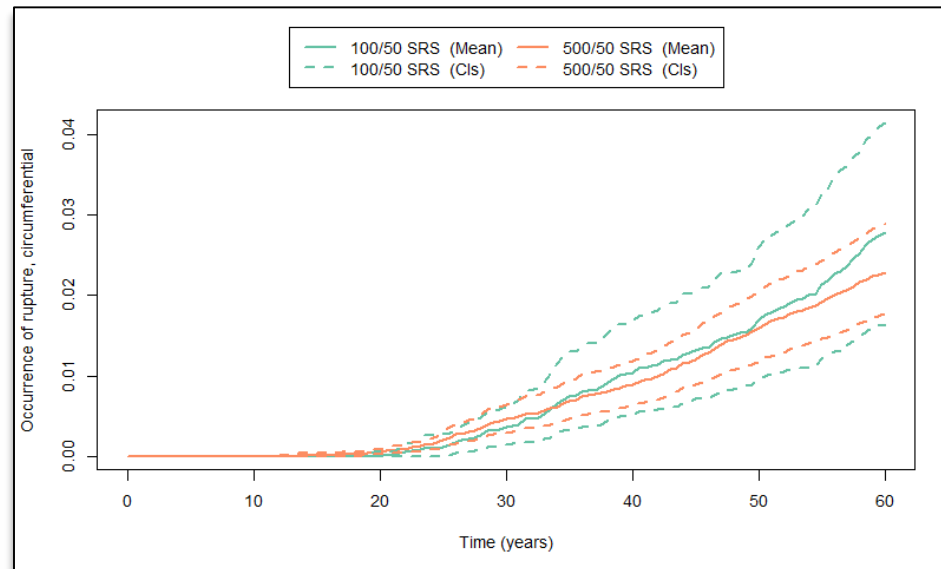
- Understand how the uncertainty in the inputs and in the model affects uncertainty in the outputs



Example: CDFs for occurrence of circ. rupture for each epistemic realization (grey), the mean (red), and the 0.05 (blue), 0.5 (green), and 0.95 (purple) quantiles.



- **Stability/convergence analysis:** Increase confidence in the results and assess their stability
 - Show that changes to the characterization of input uncertainty, selection of sampling options, and numerical accuracy is sufficient to achieve converged results



Example: Mean probability of occurrence of circ. rupture (solid line) and 95% confidence bounds (dashed lines) under two different sampling options

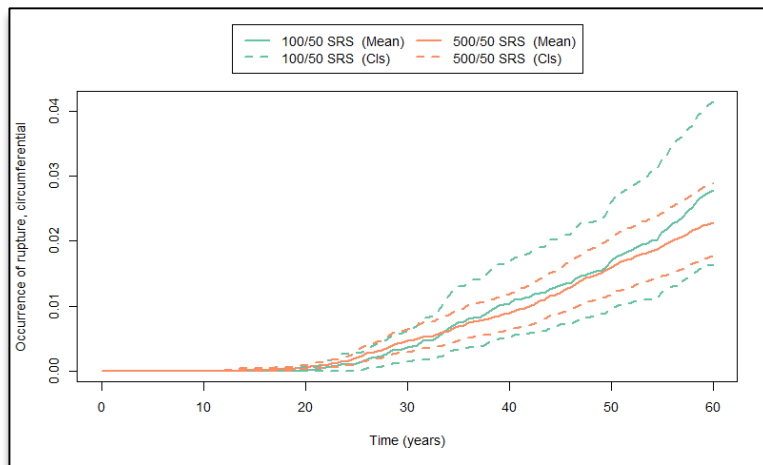


- **Sensitivity studies:** “What if?” analyses performed for individual sub- or full models
 - Study the effect of analysis assumptions including separation of uncertainty, distribution specification, and identification of problem drivers
 - Determine impact of different models on the response
 - Examples include constant vs. uncertain weld wall thickness, degree of uncertainty in PWSCC-initiated flaw size, maximum constant vs. uncertain operating stresses, etc.
 - Study alternate scenarios including worst-case scenarios, intervention scenarios, etc.

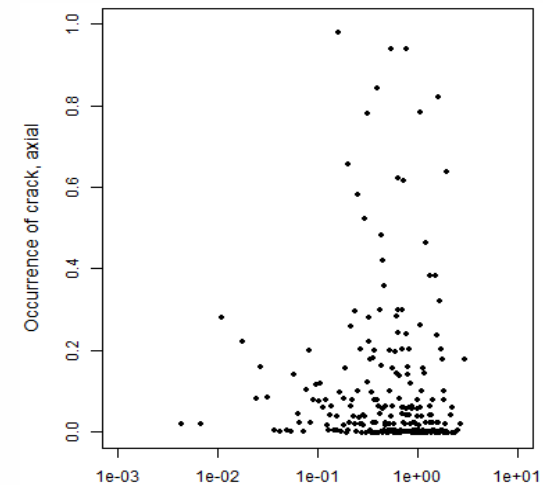


POST-PROCESSING ANALYSIS: WORKFLOW

1. Define a Quantity of Interest (QoI) or a set of QoIs
2. Determine sensitivity of the outputs to the inputs
3. Characterize sampling uncertainty of the QoI



Occurrence of circ. rupture

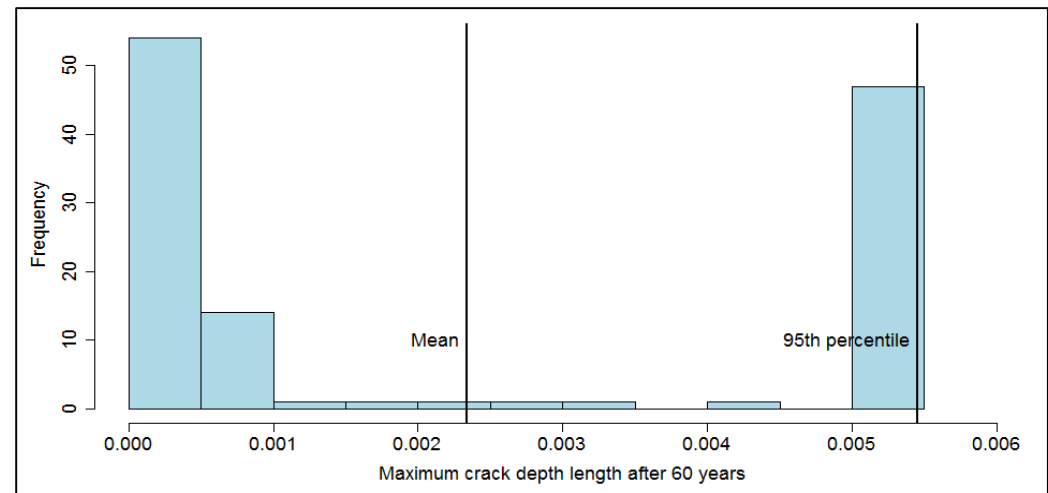


ID Hoop WRS pre-mitigation



- Multiple outputs are available in xLPR v2.0 as Qol:
 - Occurrence of rupture, occurrence of leak, crack length, etc...
- Qols are defined relative to looping:
 - For the binary output 'occurrence of rupture,' a natural Qol is the probability of rupture
 - Aleatory only (risk): probability of rupture over n_a aleatory samples for each epistemic sample
 - Aleatory and epistemic uncertainty: average probability of rupture over all $n_a * n_e$ samples

For the model output crack depth, the Qols 'mean crack depth' and '95th percentile of crack depth' can be calculated from the output distribution, represented here using a frequency histogram

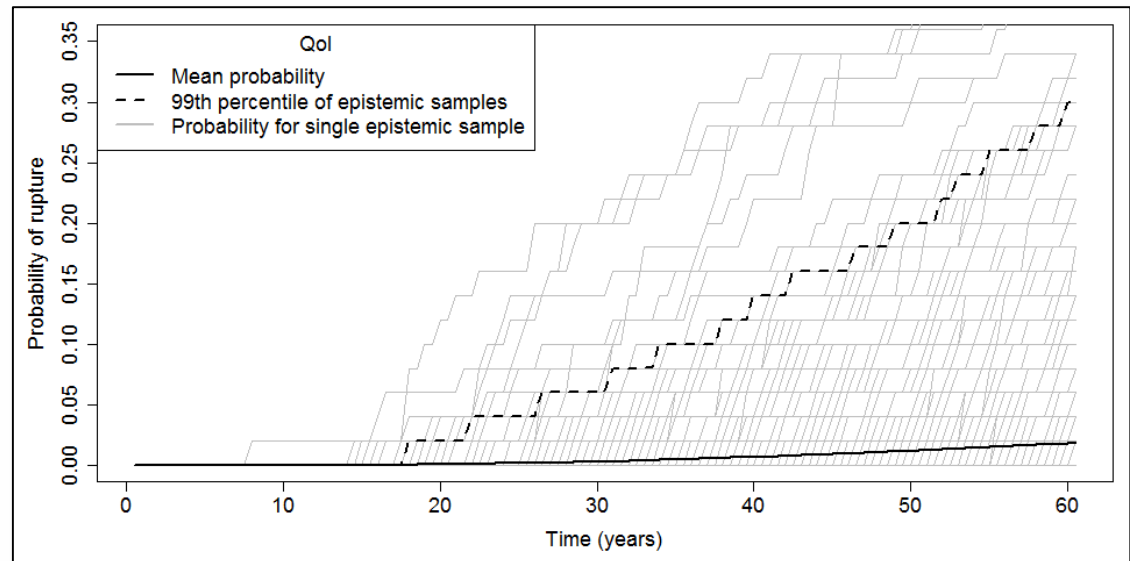




ANALYSIS OF QoI AS A FUNCTION OF SAMPLING DATA CONSIDERED

- **Best estimate:**
Mean probability over all samples (aleatory AND epistemic)
- **Knowledge of uncertainty bound:**
99th percentile of epistemic samples
- **Knowledge of variability:**
Probability for each epistemic sample

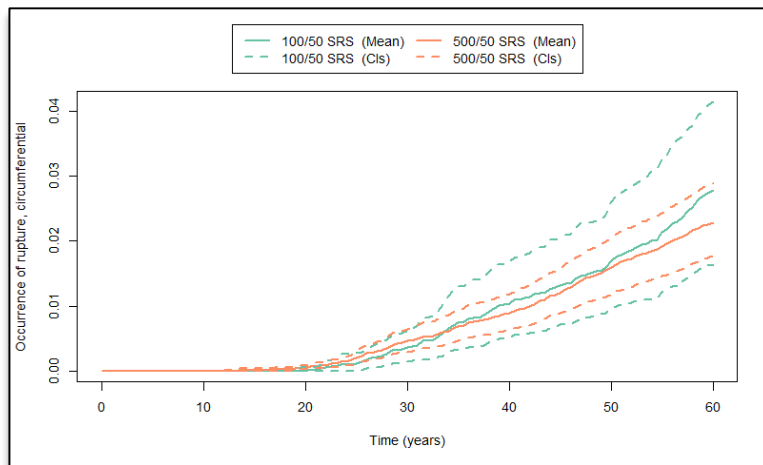
Mean probability of rupture stays below 5% at 60 years but there is a substantial amount of knowledge uncertainty in the probability of rupture as illustrated by the 99th percentile shown in the figure



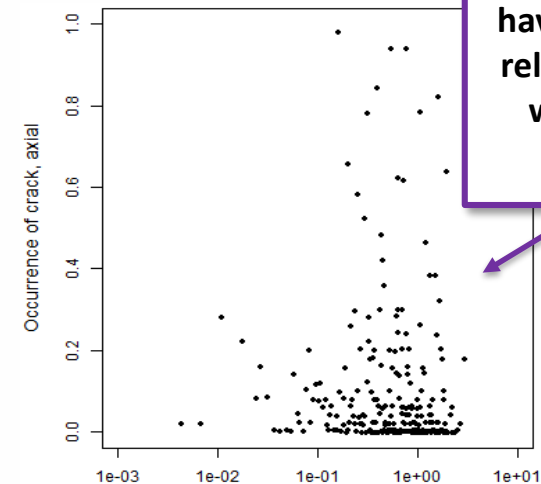


POST-PROCESSING ANALYSIS: WORKFLOW

1. Define a Quantity of Interest (QoI) or a set of QoIs
2. **Determine sensitivity of the outputs to the inputs**
3. Characterize sampling uncertainty of the QoI



Occurrence of circ. rupture

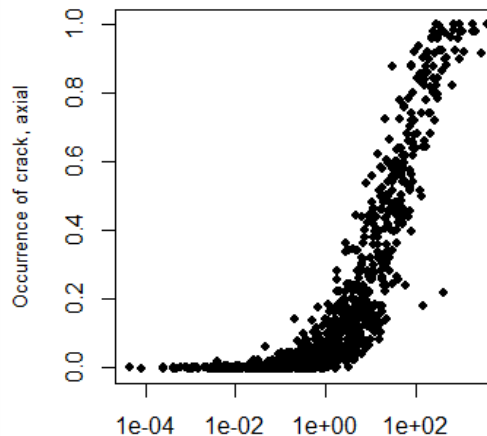


This input
appears to
have a slight
relationship
with the
output

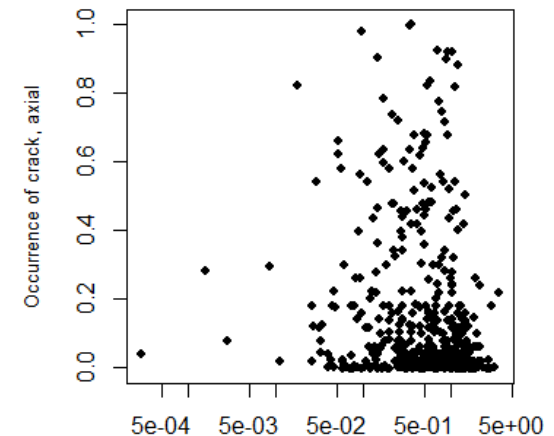


- **Sensitivity analysis (SA)** is used to:
 - Understand the **relationship between** the model **inputs and outputs**
 - Identify the inputs that have the **most significant impact** on the results of the model
- **Knowledge of the most important inputs** can be used to:
 - **Target** inputs where more information could be collected to **decrease uncertainty**
 - Identify inputs for **importance sampling** to increase precision in estimating rare probabilities

In xLPR, the DM1 multiplier is highly correlated with the probability of crack (left), while the hoop WRS pre-mitigation is not as highly correlated with the probability of crack (right)



p2543 Multiplier proport. Const. A (DM1)

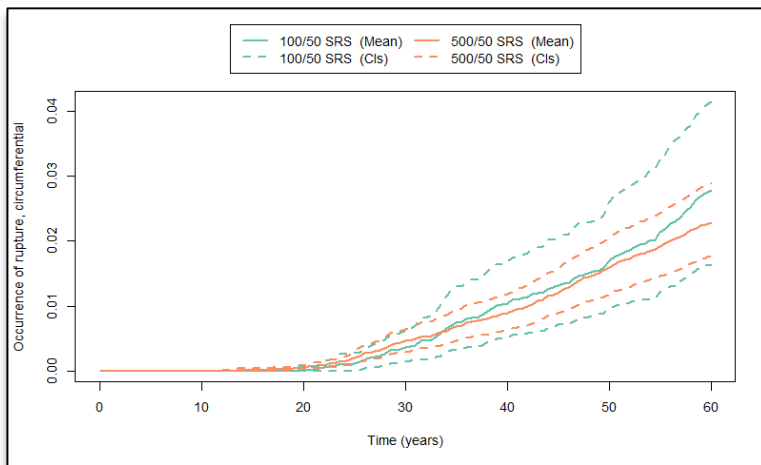


p4350 Hoop WRS Pre-mitigation

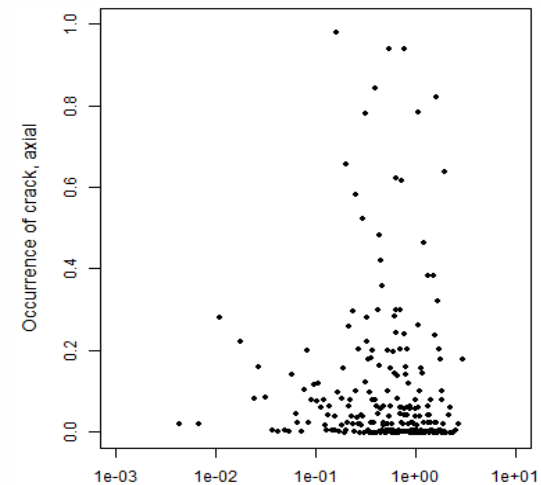


POST-PROCESSING ANALYSIS: WORKFLOW

1. Define a Quantity of Interest (QoI) or a set of QoIs
2. Determine sensitivity of the outputs to the inputs
- 3. Characterize sampling uncertainty of the QoI**



Occurrence of circ. rupture



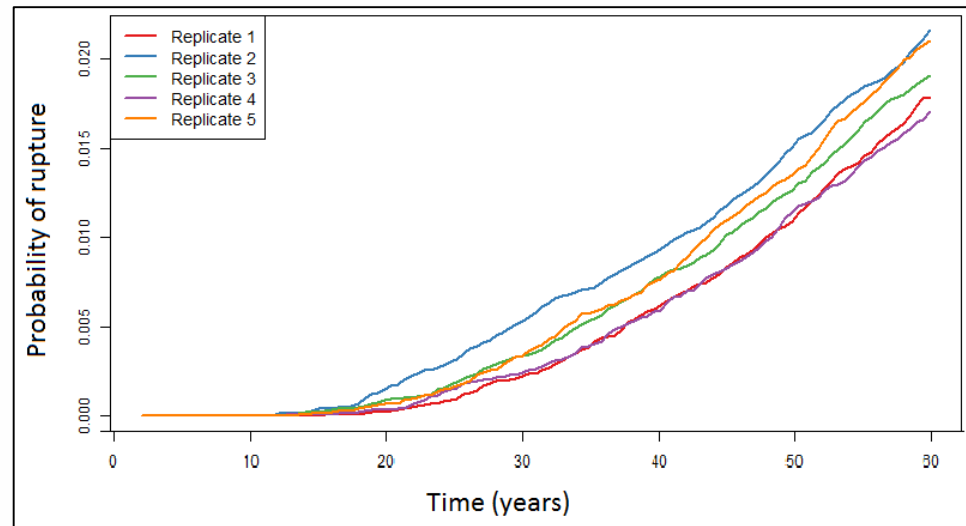
ID Hoop WRS pre-mitigation



- Reminder: The number of Monte Carlo samples is finite
- Statistical analysis methods can be used to quantify sampling uncertainty
- Running the model with different random sequences can help us understand and quantify this uncertainty
 - Accomplished by changing the random seeds in xLPR

Running a model 5 different times results in 5 different QoI estimates

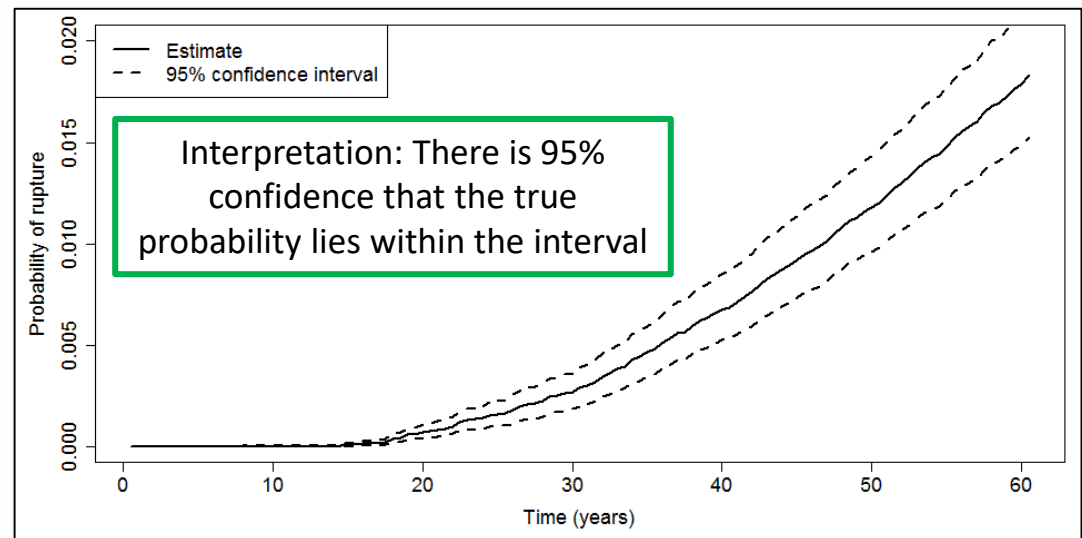
Differences in the estimates are due to sampling uncertainty



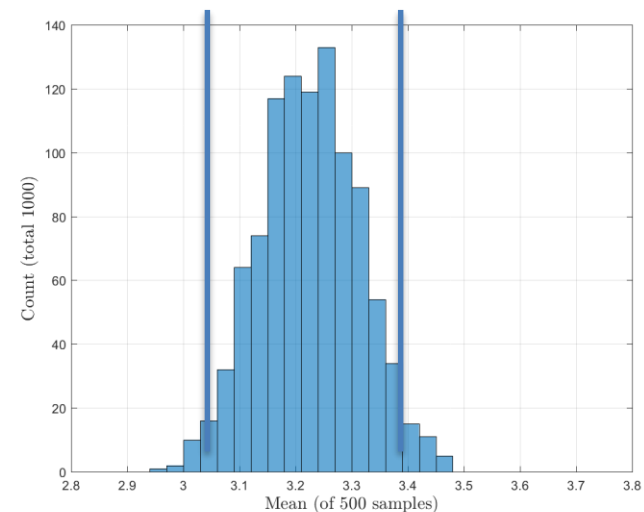
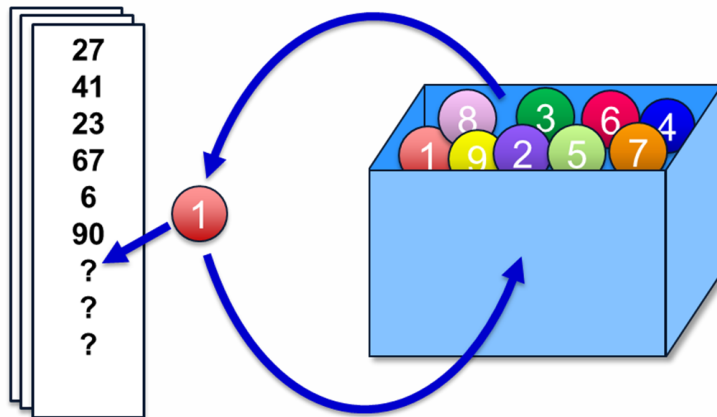


- Confidence intervals (CIs) are a common measure of sampling uncertainty
 - Bootstrap method is one of many ways to calculate confidence intervals

Sampling uncertainty in a rupture probability can be characterized using confidence intervals as an alternative to examining variability across multiple runs



- The statistical bootstrap method is one of many ways to construct CIs
 - Simple to implement, flexible, and generic statistical method
 - Observed sample (of size n) is an estimate of the population
 - RE-assemble the data (with replacement) repeatedly – each time computing the QoI
 - Collection of bootstrapped QoIs estimates sampling variability in the QoI



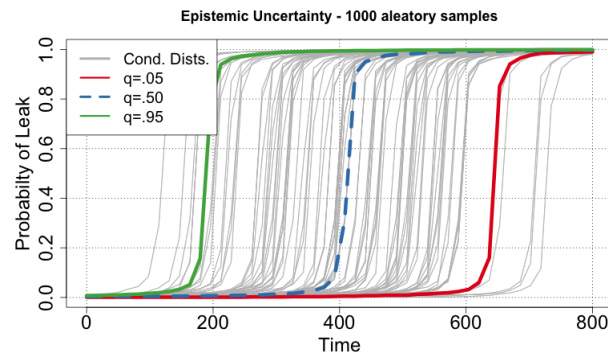
Bootstrap distribution of means (histogram) and CI (vertical lines)



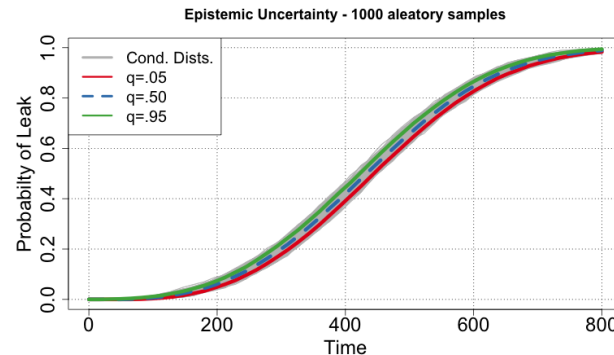
LOOPING AND ANALYSIS

- Looping over aleatory and epistemic uncertainty should be considered in the post-processing analysis
- Looping allows for uncertainty separation but can increase sampling uncertainty of 'best-estimate' Qols
 - Larger sample size may be needed to estimate small probability under looping

Example

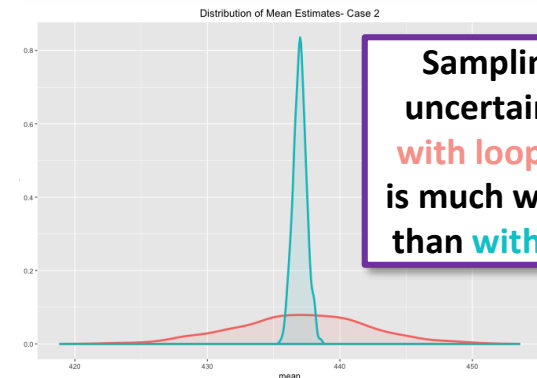


Epistemic uncertainty dominates



Aleatory uncertainty dominates

Separation of aleatory and epistemic uncertainties



Sampling
uncertainty
with looping
is much wider
than without

'Best-estimate' QoI can have much larger sampling variability when looping. Figure shows true sampling uncertainty of a best-estimate under looping (red) and no looping (blue)

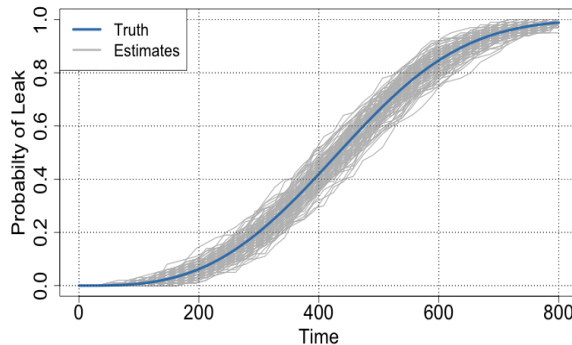


LOOPING AND ANALYSIS

Choose sample sizes so aleatory and epistemic uncertainties are not confounded by finite sample size uncertainty

Small aleatory sample size results in large sampling uncertainty when estimating a single epistemic distribution

100 estimates using 100 aleatory samples

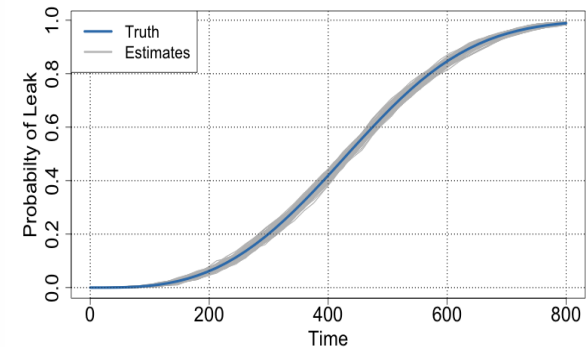


Same scenario with larger aleatory sample size



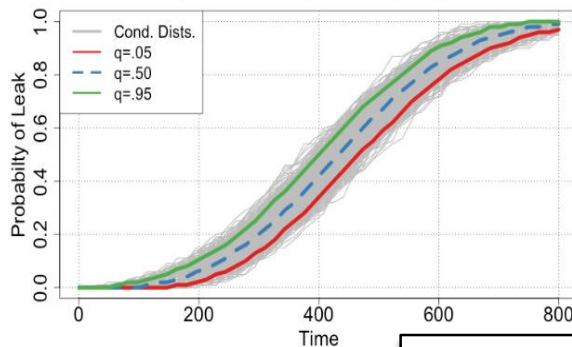
Larger aleatory sample size results in better estimation of each epistemic distribution

100 estimates using 1000 aleatory samples

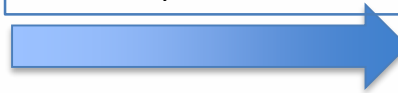


This sampling uncertainty propagates into estimation of the epistemic uncertainty

Epistemic Uncertainty - 100 aleatory samples

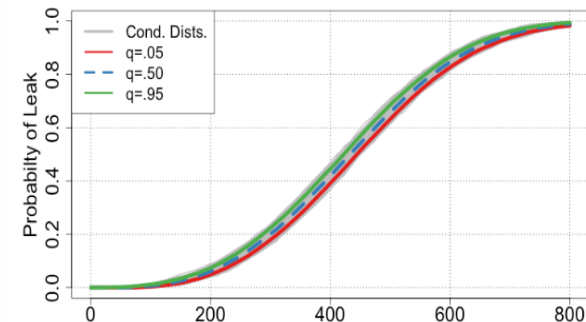


Actual epistemic uncertainty is much smaller – most of the uncertainty was sampling uncertainty



Less sampling uncertainty propagated into estimation of the epistemic uncertainty

Epistemic Uncertainty - 1000 aleatory samples



Lesson: Sample sizes must be chosen with care and are problem dependent



CONCLUSION

- **Uncertainty quantification** used for probabilistic fracture mechanics can be a **powerful tool** when applied to problems of interest
 - Can be used to determine likelihood of certain outcomes for LBB problems
- **Statistical analysis** of model results helps to refine simulations and provide the **best possible picture of risk**
 - **Uncertainty analysis** tells us how uncertainty in inputs impacts uncertainty in outputs and helps us **quantify this uncertainty**
 - **Convergence analysis** tells us whether or not our simulations are giving us a **stable result**
 - **Sensitivity analysis** can help **pinpoint inputs** for which additional study or data may be **most important**



- **Students should now be able to:**
 - Understand and compare deterministic and probabilistic fracture mechanics approaches
 - Understand the fundamental building blocks of a probabilistic fracture mechanics (PFM) analysis including:
 - Characterization of uncertainty
 - Separation of uncertainty
 - Methods for sampling inputs for Monte Carlo analysis
 - Interpretation of PFM results



What is ahead for us for the remainder of this training?



SCHEDULE

Day 1: Tuesday March 27, 2018

Time	Topic	Presenter
08:00AM - 08:30AM	Introduction and Opening Remarks	Matt Homiack (NRC/RES) Aubrey Eckert-Gallup (SNL)
08:30AM – 9:45AM	Module 1: PFM Background	Aubrey Eckert-Gallup (SNL)
09:45AM - 10:00AM	BREAK	
10:00AM - 11:30AM	Module 2: Introduction to R	Nevin Martin (SNL)
11:30AM - 12:30PM	LUNCH BREAK	
12:30PM - 01:15PM	Module 2: Continued Introduction to R	Nevin Martin (SNL)
01:15PM - 01:30PM	BREAK	
01:30PM - 04:00PM	Module 3: Sensitivity Analysis	Dusty Brooks (SNL)