DAKOTA 101

Uncertainty Quantification

http://dakota.sandia.gov

Learning goals:

- **Define uncertainty quantification and know when and why to apply it**
- **Brief survey of core UQ methods:**
  - sampling, reliability, stochastic expansions, epistemic/mixed UQ
- **Run simple example studies**
- **Understand options for post-processing the output uncertainty**

Sandia is a multi program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Risk-informed Decision Making, QMU, & UQ: An ASC V&V Program Perspective

Labs are shifting from test-based to M&S-based design and certification.

In order to support risk-informed decision-making based on M&S, we require:

- ***Predictive simulations:*** **verified and validated for application of interest**
- ***Quantified uncertainties:*** **the effect of random variability is fully understood**

DOE process: "Quantification of Margins and Uncertainties (QMU)"

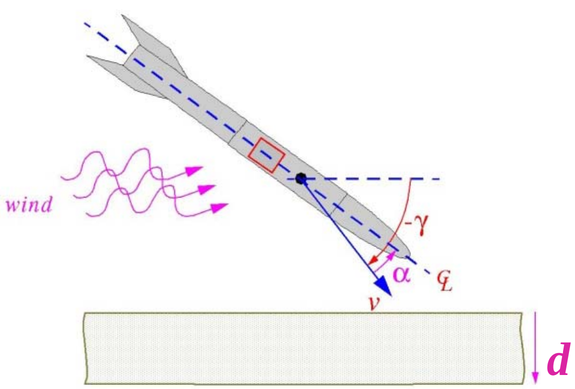
→ provide *best estimate* + *uncertainty* in the decision-making context

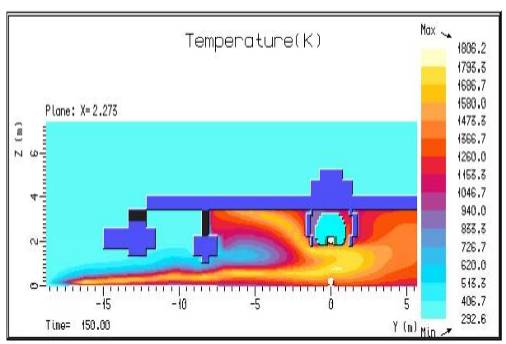
Uncertainty Quantification

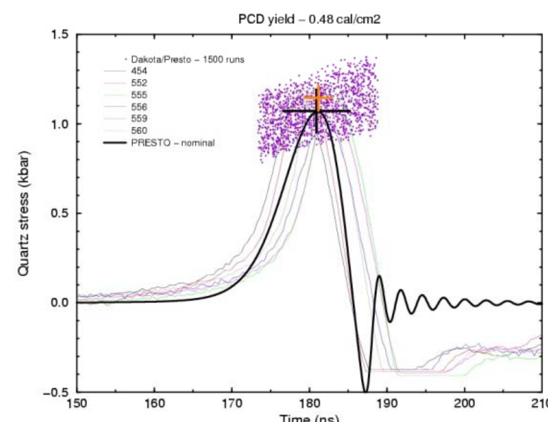
Critical component of QMU → credible M&S capability for stockpile stewardship

Uncertainty can be categorized to be one of two different types:

- **Aleatory/irreducible:** inherent variability with sufficient data → objective probabilistic models
- **Epistemic/reducible:** uncertainty from lack of knowledge → subjective probabilistic, nonprobabilistic



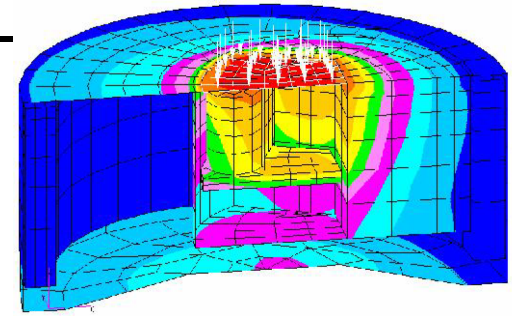


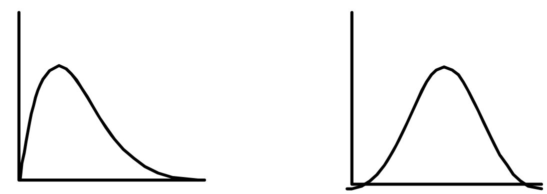


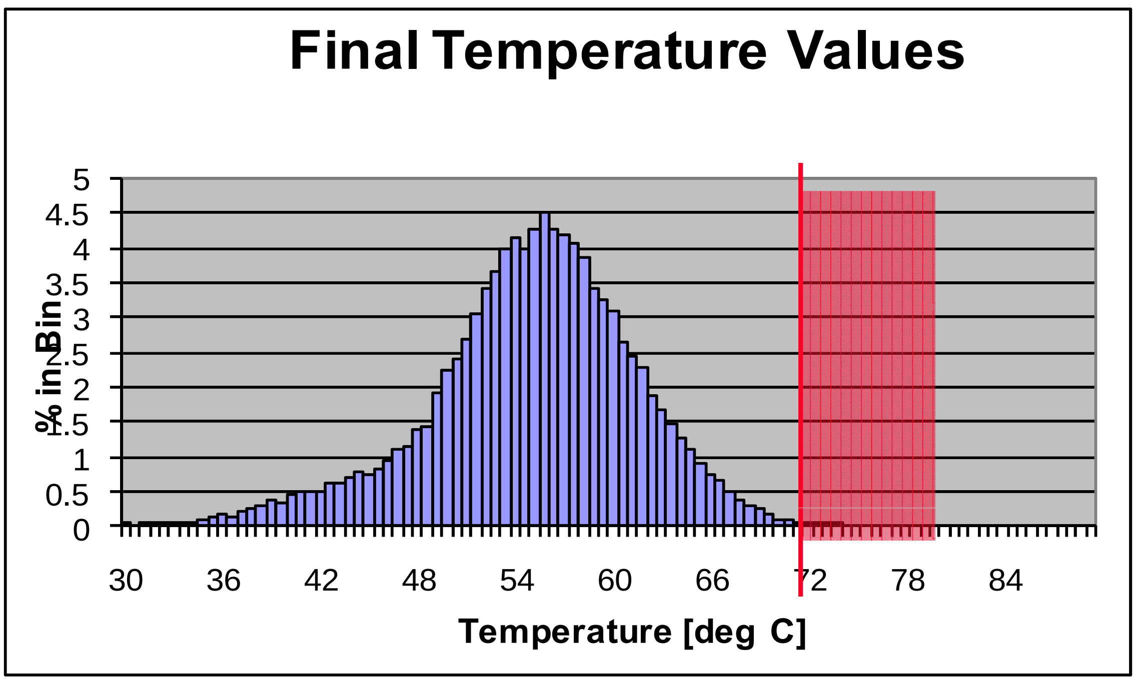
Uncertainty applications: penetration, joint mechanics, abnormal environments, shock physics, …

Example:
Thermal Uncertainty Quantification

- **Device subject to heating (experiment or computational simulation)**
- **Uncertainty in composition/ environment (thermal conductivity, density, boundary), parameterized by $u_1, \ldots, u_N$**
- **Response temperature $f(u)=T(u_1, \ldots, u_N)$ calculated by heat transfer code**







Given distributions of $u_1,\ldots,u_N$, UQ methods calculate statistical info on outputs:

- **Mean(T), StdDev(T), Probability($T \geq T_{critical}$)**
- **Probability distribution of temperatures**
- ***Correlations (trends) and sensitivity of temperature***

Uncertainty Quantification Algorithms @ SNL: New methods bridge robustness/efficiency gap

	Production	New	Under dev.	Planned	Collabs.
Sampling	Latin Hypercube, Monte Carlo	Importance, Incremental		Bootstrap, Jackknife	FSU
Reliability	*Local:* Mean Value, First-order & second-order reliability methods (FORM, SORM)	*Global:* Efficient global reliability analysis (EGRA)	gradient-enhanced	Ensemble emulator-based	*Local:* Notre Dame, *Global:* Vanderbilt
Stochastic expansion		Polynomial chaos expansion, Stochastic collocation	Dimension-adaptive p-/h-refinement, gradient-enhanced	Region-adaptive, discrete, multi-physics	Stanford, Purdue, Austr. Natl., FSU
Other probabilistic		Random fields/ stochastic proc.		Dimension reduction	Cornell, Maryland
Epistemic	Interval-valued/ Second-order prob. (nested sampling)	Opt-based interval estimation, Dempster-Shafer	Bayesian	Imprecise probability	LANL, Applied Biometrics
Metrics & Global SA	Importance factors, Partial correlations	Main effects, Variance-based decomposition	Stepwise regression		UNM

Reseach: Tailoring & Adaptivity

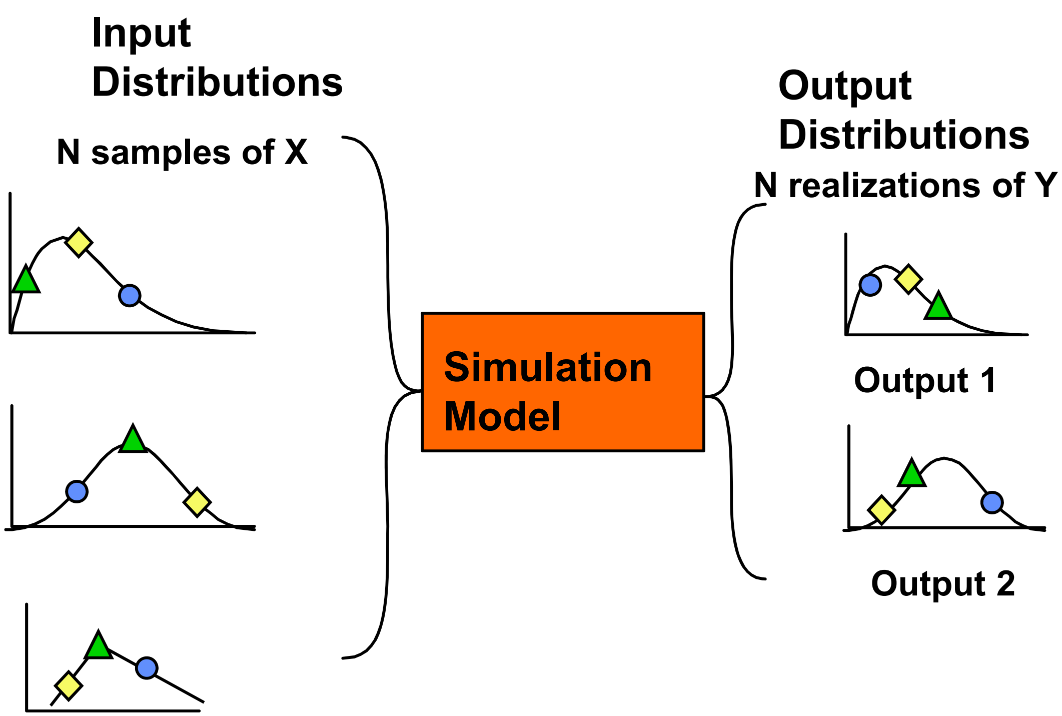
Adv. Deployment

Fills Gaps

Sampling Methods

Common UQ Method: Random Sampling

- Assume distributions on each of the *n* uncertain input variables
- Sample from each distribution and pair into *N* samples
- Run the simulation model for each of the *N* samples
- Use results ensemble to build up a distribution for each of the *m* outputs



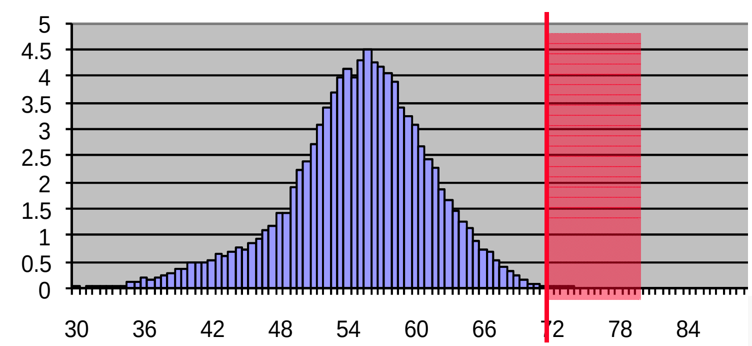
- sample mean

$$\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T(u^i)$$

- sample variance

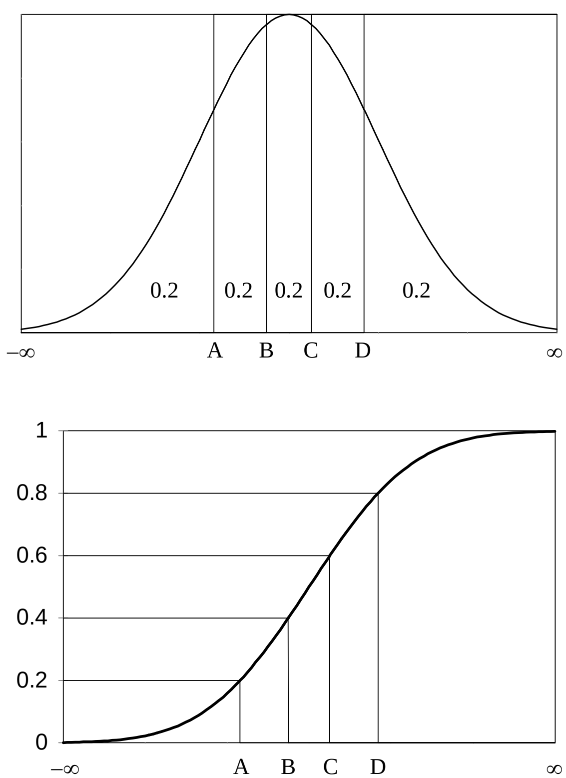
$$T_{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N}\left[T(u^i) - \bar{T}\right]^2$$

- full PDF(probabilities)



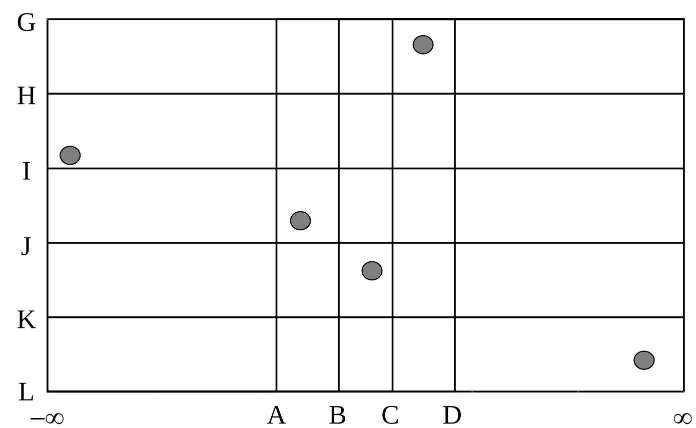
Latin Hypercube Sampling

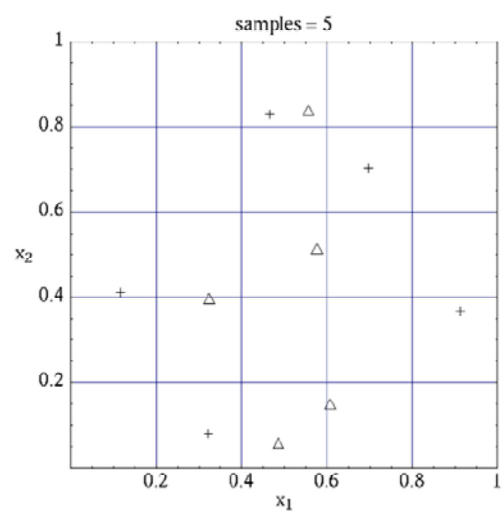
- **LHS is stratified random sampling among equal probability bins for all 1-D projections of an n-dimensional set of samples.**
  - **Early work by McKay and Conover**
  - **Restricted pairing by Iman → enforce prescribed input correlations**

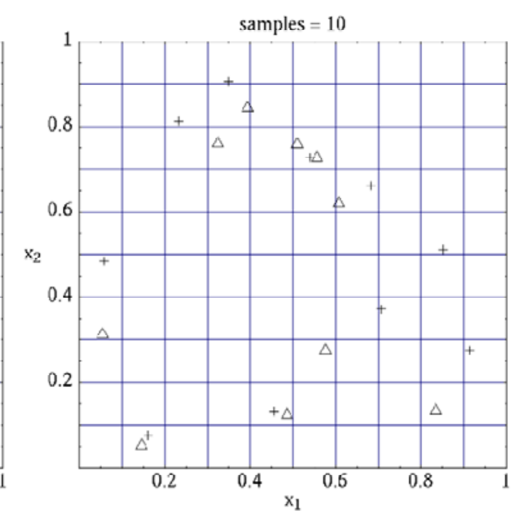


Intervals Used with a LHS of Size N = 5 in Terms of the PDF and CDF for a Normal Random Variable

A possible LHS for n=2, N=5 with X1 = normal and X2 = uniform



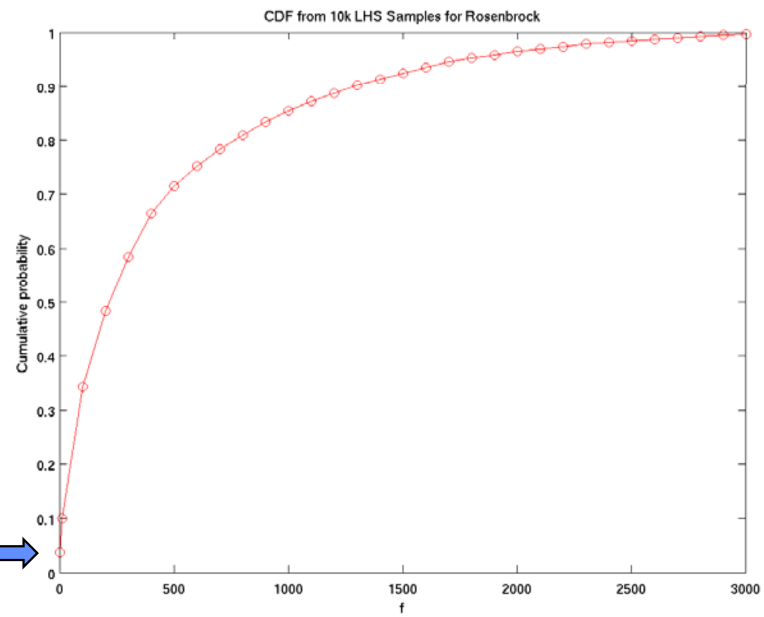


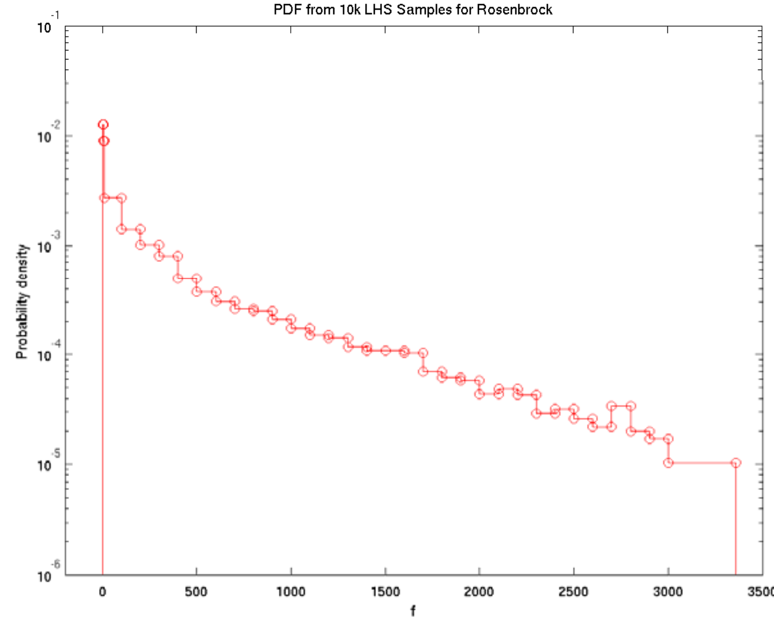


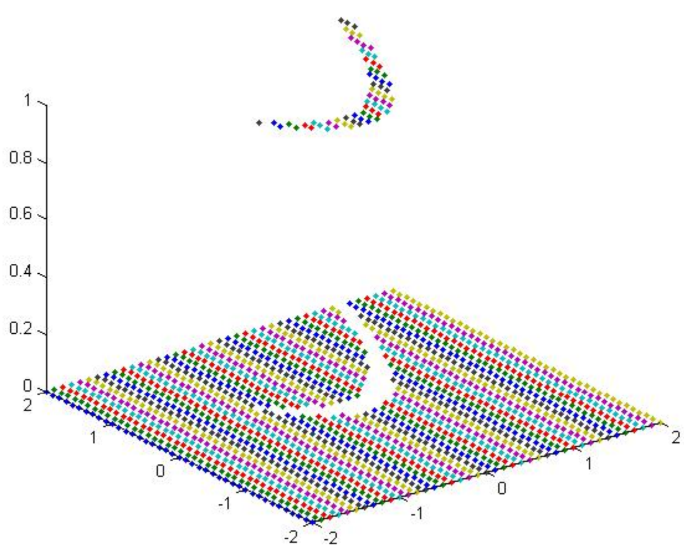
Sandia National Laboratories

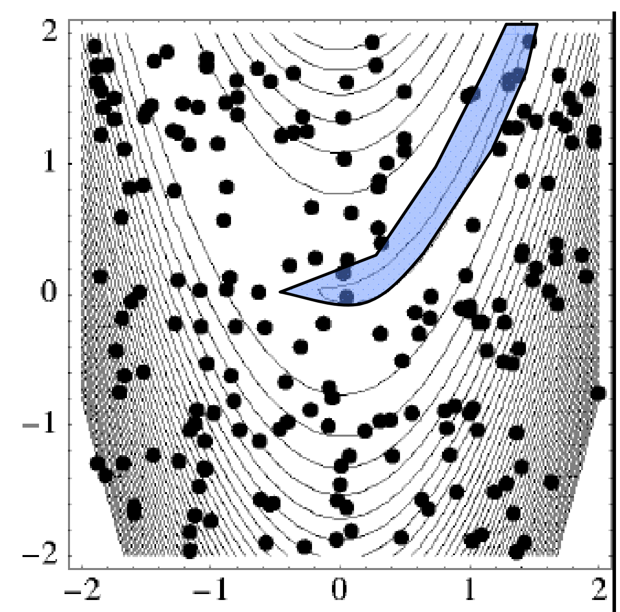
Example: Probability(rosenbrock < 3)

p(f<3)=.0376









Class Exercise: Cantilever Beam UQ with Sampling

- **Perform UQ with LHS method on `mod_cantilever`**
- **Determine mean system response, variability, margin to failure given (see variables section of reference manual)**
  - **Yield stress** — **R ~ Normal(40000, 2000)**
  - **Young's modulus** — **E ~ Normal(2.9e7, 1.45e6)**
  - **Horizontal load** — **X ~ Normal(500, 100)**
  - **Vertical load** — **Y ~ Normal(1000, 100)**
- **Hold width and thickness at 2.5**
- **Use `probability_levels` or `response_levels` in method**
- **Extra exercises (time permitting)**
  - ***What happens to confidence intervals on the mean and standard deviation as number of samples varies?***
  - **Instead of normal, try uniform distribution for each random variable. What do you expect would happen?**

Example Input/Output: Sampling

extraexamples/dakota uq cantilever lhs.in

```
strategy
       single_method graphics

method,
      sampling
        sample_type lhs
        samples = 10000   seed = 12347
        num_probability_levels = 0 17 17
        probability_levels =
        .001 .01 .05 .1 .15 .2 .3 .4 .5 .6 .7 .8 .85 .9 .95 .99 .999
        .001 .01 .05 .1 .15 .2 .3 .4 .5 .6 .7 .8 .85 .9 .95 .99 .999
        cumulative distribution
#      output silent

variables,
      continuous_design = 2
         initial_point 2.5 2.5
         upper_bounds 10.0 10.0
         lower_bounds 1.0 1.0
         descriptors 'beam_width' 'beam_thickness'
      normal_uncertain = 4
         means = 40000. 29.E+6 500. 1000.
         std_deviations = 2000. 1.45E+6 100. 100.
         descriptors = 'R' 'E' 'X' 'Y'

interface,
      direct analysis_driver = 'mod_cantilever'

responses,
      num_response_functions = 3
      no_gradients
      no_hessians
```

Input (extra_examples/dakota_uq_cantilever_lhs.in)

```
Statistics based on 10000 samples:

Moment-based statistics for each response function:
                            Mean           Std Dev          Skewness          Kurtosis
       area  6.2500000000e+00  0.0000000000e+00              -nan              -nan
   g_stress  1.7599759864e+04  5.7886440706e+03 -2.2153567379e-02 -4.9234550018e-02
    g_displ  1.7201261575e+00  4.0670385498e-01  1.7796424852e-01  8.0009704624e-02

95% confidence intervals for each response function:
                     LowerCI_Mean      UpperCI_Mean    LowerCI_StdDev    UpperCI_StdDev
       area  6.2500000000e+00  6.2500000000e+00  0.0000000000e+00  0.0000000000e+00
   g_stress  1.7486290789e+04  1.7713228938e+04  5.7095204696e+03  5.8700072185e+03
    g_displ  1.7121539434e+00  1.7280983716e+00  4.0114471657e-01  4.1242034152e-01

Level mappings for each response function:
Cumulative Distribution Function (CDF) for g_stress:
     Response Level  Probability Level  Reliability Index  General Rel Index
     --------------  -----------------  -----------------  -----------------
   2.4921421856e+02   1.0000000000e-03
   4.1489075797e+03   1.0000000000e-02
   7.9708753041e+03   5.0000000000e-02
   1.0090342657e+04   1.0000000000e-01
   1.1589780322e+04   1.5000000000e-01
   1.2731567123e+04   2.0000000000e-01
   1.4564078343e+04   3.0000000000e-01
   1.6151010310e+04   4.0000000000e-01
   1.7689441098e+04   5.0000000000e-01
   1.9129203866e+04   6.0000000000e-01
   2.0683233939e+04   7.0000000000e-01
   2.2457356004e+04   8.0000000000e-01
   2.3589089220e+04   8.5000000000e-01
   2.4920875151e+04   9.0000000000e-01
   2.7044322788e+04   9.5000000000e-01
   3.0752664401e+04   9.9000000000e-01
   3.5331778223e+04   9.9900000000e-01
Cumulative Distribution Function (CDF) for g_displ:
     Response Level  Probability Level  Reliability Index  General Rel Index
     --------------  -----------------  -----------------  -----------------
   5.8671224313e-01   1.0000000000e-03
   8.3043198982e-01   1.0000000000e-02
   1.0603181923e+00   5.0000000000e-02
   1.2097669707e+00   1.0000000000e-01
   1.2966885568e+00   1.5000000000e-01
   1.3746930033e+00   2.0000000000e-01
   1.5000347941e+00   3.0000000000e-01
   1.6076526708e+00   4.0000000000e-01
   1.7093341472e+00   5.0000000000e-01
   1.8116530545e+00   6.0000000000e-01
   1.9242640620e+00   7.0000000000e-01
   2.0601011526e+00   8.0000000000e-01
   2.1419095361e+00   8.5000000000e-01
   2.2454142842e+00   9.0000000000e-01
   2.3966062187e+00   9.5000000000e-01
   2.7299039059e+00   9.9000000000e-01
   3.0858487193e+00   9.9900000000e-01
```

Output

Reliability Methods

UQ with Reliability Methods: Mean Value Method

Linear approximation for first two moments:

$$\begin{aligned}\mu_g &= g(\mu_{\mathbf{x}}) \\ \sigma_g^2 &= \sum_i \sum_j Cov(i,j) \frac{dg}{dx_i}(\mu_{\mathbf{x}}) \frac{dg}{dx_j}(\mu_{\mathbf{x}})\end{aligned}$$

Projection of moments for reliability indices:

$$\bar{z} \to p, \beta \begin{cases} \beta_{cdf} = \dfrac{\mu_g - \bar{z}}{\sigma_g} \\ \beta_{ccdf} = \dfrac{\bar{z} - \mu_g}{\sigma_g} \end{cases}$$

$$\bar{p}, \bar{\beta} \to z \begin{cases} z = \mu_g - \sigma_g \bar{\beta}_{cdf} \\ z = \mu_g + \sigma_g \bar{\beta}_{ccdf} \end{cases}$$

Normality assumption:

$$\begin{cases} p(g \leq z) = \Phi(-\beta_{cdf}) \\ p(g > z) = \Phi(-\beta_{ccdf}) \end{cases}$$

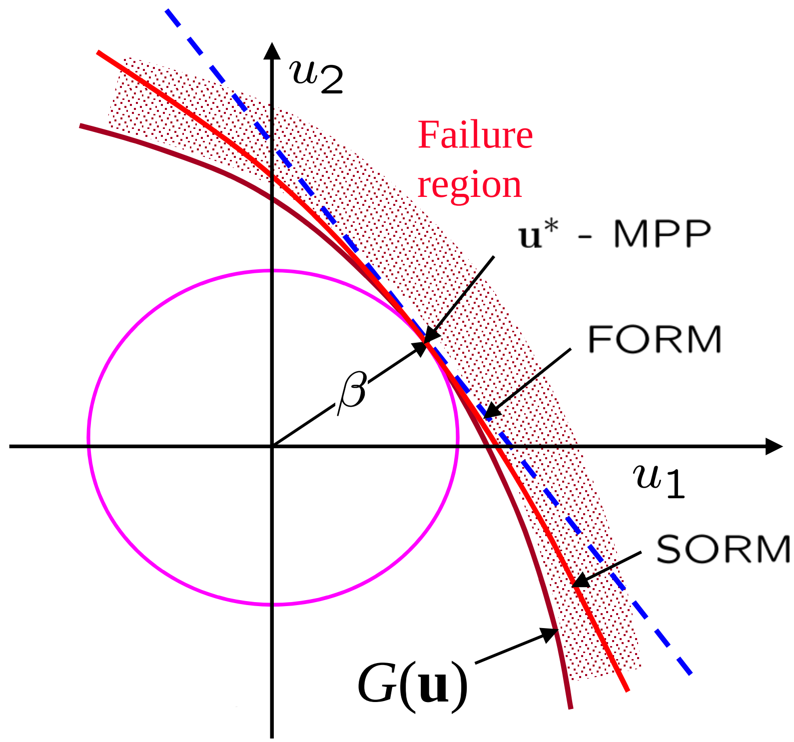
Rough statistics

UQ with Reliability Methods: MPP Search Methods

<u>Reliability Index Approach (RIA)</u>

$$\begin{aligned} &\text{minimize} \quad \mathbf{u}^T\mathbf{u} \\ &\text{subject to} \quad G(\mathbf{u}) = \bar{z} \end{aligned}$$

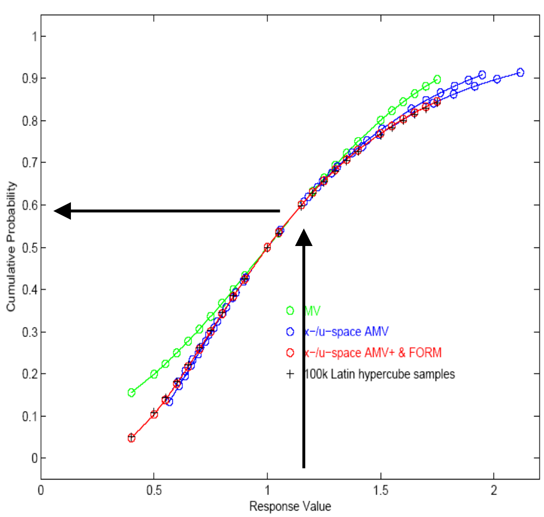
Find min dist to G level curve
Used for fwd map $z \rightarrow p/\beta$



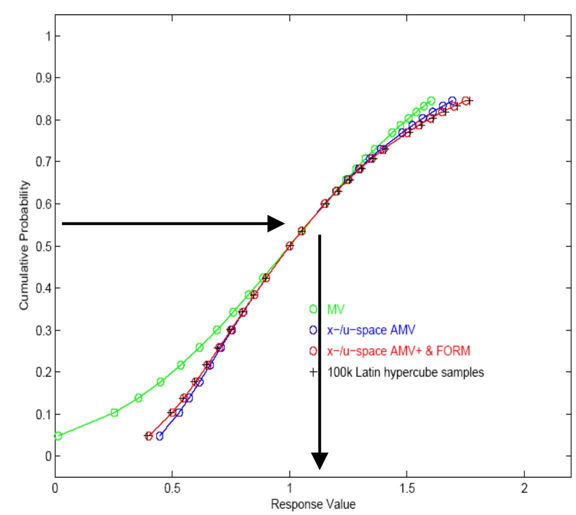
<u>Performance Measure Approach (PMA)</u>

$$\begin{aligned} &\text{minimize} \quad \pm G(\mathbf{u}) \\ &\text{subject to} \quad \mathbf{u}^T\mathbf{u} = \bar{\beta}^2 \end{aligned}$$

Find min G at β radius
Used for inv map $p/\beta \rightarrow z$



Nataf x → u: $$\begin{aligned} \Phi(z_i) &= F(x_i) \\ \mathbf{z} &= \mathbf{L}\mathbf{u} \end{aligned}$$



Local Reliability Example

- **Compare sampling methods to local reliability methods (modify `dakota_uq_cantilever_lhs.in` or create `dakota_uq_cantilever_rel.in`)**

Example Input/Output: Reliability

```
strategy
     single_method #graphics

method,
     local_reliability
       mpp_search no_approx
       num_probability_levels = 0 17 17
       probability_levels =
      .001 .01 .05 .1 .15 .2 .3 .4 .5 .6 .7 .8 .85 .9 .95 .99 .999
      .001 .01 .05 .1 .15 .2 .3 .4 .5 .6 .7 .8 .85 .9 .95 .99 .999
       cumulative distribution

variables,
     continuous_design = 2
       initial_point    2.5    2.5
       upper_bounds    10.0   10.0
       lower_bounds     1.0    1.0
       descriptors      'beam_width' 'beam_thickness'
     normal_uncertain = 4
       means          =  40000. 29.E+6 500. 1000.
       std_deviations =  2000. 1.45E+6 100. 100.
       descriptors       =  'R' 'E' 'X' 'Y'

interface,
     direct
       analysis_driver = 'mod_cantilever'

responses,
     descriptors = 'area' 'g_stress' 'g_displ'
     num_response_functions = 3
     analytic_gradients
     no_hessians
```

```
<<<<< Function evaluation summary: 437 total (435 new, 2 duplicate)
-----------------------------------------------------------------
Cumulative Distribution Function (CDF) for g_stress:
     Response Level  Probability Level  Reliability Index  General Rel Index
     --------------  -----------------  -----------------  -----------------
  -2.8366322206e+02   9.9999999999e-04   3.0902323062e+00   3.0902323062e+00
   4.1370568311e+03   1.0000000000e-02   2.3263478740e+00   2.3263478740e+00
   8.0809718495e+03   5.0000000000e-02   1.6448536270e+00   1.6448536270e+00
   1.0183458352e+04   1.0000000000e-01   1.2815515655e+00   1.2815515655e+00
   1.1601996014e+04   1.5000000000e-01   1.0364333895e+00   1.0364333895e+00
   1.2729404779e+04   2.0000000000e-01   8.4162123357e-01   8.4162123357e-01
   1.4565211274e+04   3.0000000000e-01   5.2440051271e-01   5.2440051271e-01
   1.6133840235e+04   4.0000000000e-01   2.5334710314e-01   2.5334710314e-01
   1.7599296043e+04   4.9995140020e-01   1.2182163305e-04   1.2182163305e-04
   1.9066159765e+04   6.0000000000e-01  -2.5334710314e-01  -2.5334710314e-01
   2.0634788726e+04   7.0000000000e-01  -5.2440051271e-01  -5.2440051271e-01
   2.2470595221e+04   8.0000000000e-01  -8.4162123357e-01  -8.4162123357e-01
   2.3598003986e+04   8.5000000000e-01  -1.0364333895e+00  -1.0364333895e+00
   2.5016541648e+04   9.0000000000e-01  -1.2815515655e+00  -1.2815515655e+00
   2.7119028150e+04   9.5000000000e-01  -1.6448536270e+00  -1.6448536270e+00
   3.1062943169e+04   9.9000000000e-01  -2.3263478740e+00  -2.3263478740e+00
   3.5483663222e+04   9.9900000000e-01  -3.0902323062e+00  -3.0902323062e+00
Cumulative Distribution Function (CDF) for g_displ:
     Response Level  Probability Level  Reliability Index  General Rel Index
     --------------  -----------------  -----------------  -----------------
   5.2190342194e-01   1.0000000000e-03   3.0902323062e+00   3.0902323062e+00
   7.9940521862e-01   1.0000000000e-02   2.3263478740e+00   2.3263478740e+00
   1.0530040474e+00   5.0000000000e-02   1.6448536270e+00   1.6448536270e+00
   1.1908432559e+00   1.0000000000e-01   1.2815515655e+00   1.2815515655e+00
   1.2849790018e+00   1.5000000000e-01   1.0364333895e+00   1.0364333895e+00
   1.3604842916e+00   2.0000000000e-01   8.4162123357e-01   8.4162123357e-01
   1.4848060107e+00   3.0000000000e-01   5.2440051271e-01   5.2440051271e-01
   1.5924532498e+00   4.0000000000e-01   2.5334710314e-01   2.5334710314e-01
   1.6943072241e+00   4.9999077279e-01   2.3129179837e-05   2.3129179837e-05
   1.7974470361e+00   6.0000000000e-01  -2.5334710314e-01  -2.5334710314e-01
   1.9092590248e+00   7.0000000000e-01  -5.2440051271e-01  -5.2440051271e-01
   2.0421553573e+00   8.0000000000e-01  -8.4162123357e-01  -8.4162123357e-01
   2.1249148356e+00   8.5000000000e-01  -1.0364333895e+00  -1.0364333895e+00
   2.2303425020e+00   9.0000000000e-01  -1.2815515655e+00  -1.2815515655e+00
   2.3893977168e+00   9.5000000000e-01  -1.6448536270e+00  -1.6448536270e+00
   2.6975188966e+00   9.9000000000e-01  -2.3263478740e+00  -2.3263478740e+00
   3.0598390500e+00   9.9900000000e-01  -3.0902323062e+00  -3.0902323062e+00
```

Reliability Input (examples/methods/dakota_uq_reliability.in)

Stochastic Expansion Methods

Polynomial Chaos Expansions

Approximate response w/ spectral proj. using orthogonal polynomial basis fns

i.e. $\boxed{R = \sum_{j=0}^{P} \alpha_j \Psi_j(\boldsymbol{\xi})}$ using

$$\begin{array}{lclcl}
\Psi_0(\boldsymbol{\xi}) & = & \psi_0(\xi_1)\,\psi_0(\xi_2) & = & 1 \\
\Psi_1(\boldsymbol{\xi}) & = & \psi_1(\xi_1)\,\psi_0(\xi_2) & = & \xi_1 \\
\Psi_2(\boldsymbol{\xi}) & = & \psi_0(\xi_1)\,\psi_1(\xi_2) & = & \xi_2 \\
\Psi_3(\boldsymbol{\xi}) & = & \psi_2(\xi_1)\,\psi_0(\xi_2) & = & \xi_1^2 - 1 \\
\Psi_4(\boldsymbol{\xi}) & = & \psi_1(\xi_1)\,\psi_1(\xi_2) & = & \xi_1\xi_2 \\
\Psi_5(\boldsymbol{\xi}) & = & \psi_0(\xi_1)\,\psi_2(\xi_2) & = & \xi_2^2 - 1
\end{array}$$

- Nonintrusive: estimate α_j using sampling, regression, tensor-product quadrature, sparse grids, or cubature

$$\boxed{\alpha_j = \frac{\langle R, \Psi_j \rangle}{\langle \Psi_j^2 \rangle} = \frac{1}{\langle \Psi_j^2 \rangle} \int_\Omega R\,\Psi_j\,\varrho(\boldsymbol{\xi})\,d\boldsymbol{\xi}}$$

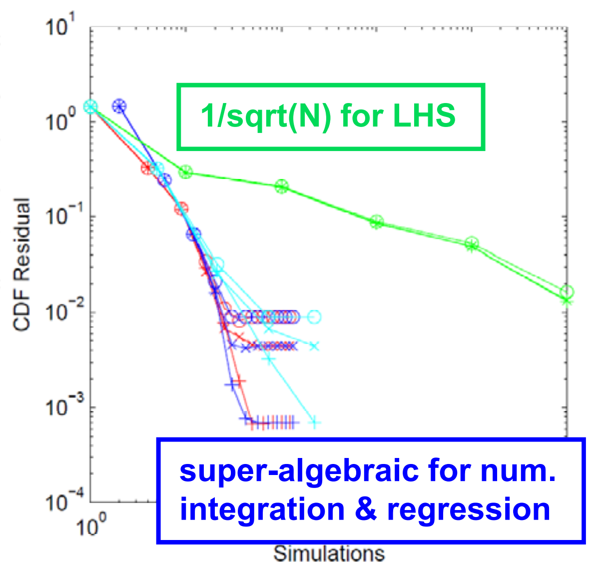
$$\boxed{\langle \Psi_j^2 \rangle = \prod_{i=1}^{n} \langle \psi_{m_i^j}^2 \rangle}$$

Generalized PCE (Wiener-Askey + numerically-generated)

- Tailor basis: selection of basis orthogonal to input PDF avoids additional nonlinearity

Distribution	Density function	Polynomial	Weight function	Support range
Normal	$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$	Hermite $He_n(x)$	$e^{\frac{-x^2}{2}}$	$[-\infty, \infty]$
Uniform	$\frac{1}{2}$	Legendre $P_n(x)$	1	$[-1, 1]$
Beta	$\frac{(1-x)^\alpha (1+x)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)}$	Jacobi $P_n^{(\alpha,\beta)}(x)$	$(1-x)^\alpha (1+x)^\beta$	$[-1, 1]$
Exponential	e^{-x}	Laguerre $L_n(x)$	e^{-x}	$[0, \infty]$
Gamma	$\frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}$	Generalized Laguerre $L_n^{(\alpha)}(x)$	$x^\alpha e^{-x}$	$[0, \infty]$

Additional bases generated numerically (discretized Stieltjes + Golub-Welsch)



- Tailor expansion form:
  – Dimension p-refinement: anisotropic TPQ/SSG, generalized SSG
  – Dimension & region h-refinement: local bases with global & local refinement

Stochastic Collocation
(based on interpolation polynomials)

Instead of estimating coefficients for known basis functions, form <u>interpolants</u> for known coefficients

$$R = \sum_{j=1}^{N_p} r_j \boldsymbol{L}_j(\boldsymbol{\xi})$$

- **Lagrange (values) or Hermite (values+derivatives)**
- **Global (high order) or local (low order spline)**

$$L_i = \prod_{\substack{j=1 \\ j \neq i}}^{m} \frac{x - x_j}{x_i - x_j}$$

$$R(\boldsymbol{\xi}) \cong \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} r\left(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}\right) \left(L_{j_1}^{i_1} \otimes \cdots \otimes L_{j_n}^{i_n}\right)$$

Sparse interpolants formed using Σ of tensor interpolants

Advantages relative to PCE:

- Somewhat simpler (no expansion order to manage separately)
- Often less expensive (no integration for coefficients)
- Expansion only formed for sampling → probabilities (estimating moments of any order is straightforward)
- Adaptive h-refinement with hierarchical surpluses; explicit gradient-enhancement

Disadvantages relative to PCE:

- Less flexible/fault tolerant → structured data sets (tensor/sparse grids)
- Expansion variance not guaranteed positive (important in opt./interval est.)
- No direct inference of spectral decay rates

With sufficient care on PCE form, PCE/SC performance is essentially identical for many cases of interest (tensor/sparse grids with standard Gauss rules)

Approaches for forming PCE/SC Expansions

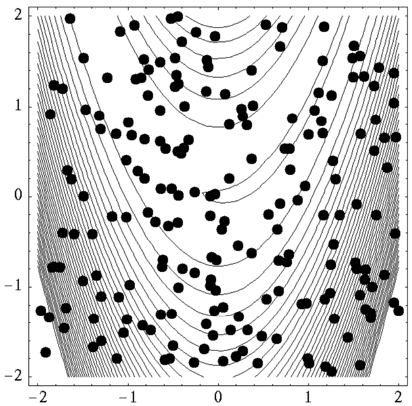
Random sampling: PCE

Expectation (sampling):

- Sample w/i distribution of ξ
- Compute expected value of product of $\boldsymbol{R}$ and each Ψ_j

Linear regression ("point collocation"):

- Sample w/i distribution of ξ
- Solves least squares data fit for all coefficients at once:



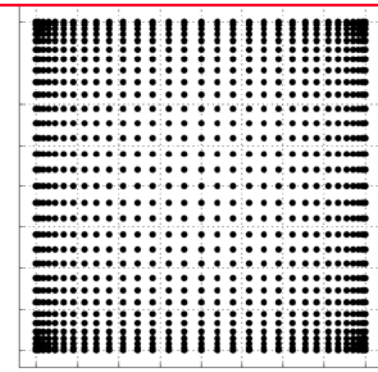
$$\boxed{\Psi\alpha = R}$$

Tensor-product quadrature: PCE/SC

$$\boxed{\mathscr{U}^i(f)(\xi) = \sum_{j=1}^{m_i} f(\xi^i_j)\, w^i_j}$$

$$\boxed{Q^n_{\mathbf{i}} f(\xi) = (\mathscr{U}^{i_1}\otimes\cdots\otimes\mathscr{U}^{i_n})(f)(\xi) = \sum_{j_1=1}^{m_{i_1}}\cdots\sum_{j_n=1}^{m_{i_n}} f\left(\xi^{i_1}_{j_1},\dots,\xi^{i_n}_{j_n}\right)\left(w^{i_1}_{j_1}\otimes\cdots\otimes w^{i_n}_{j_n}\right)}$$

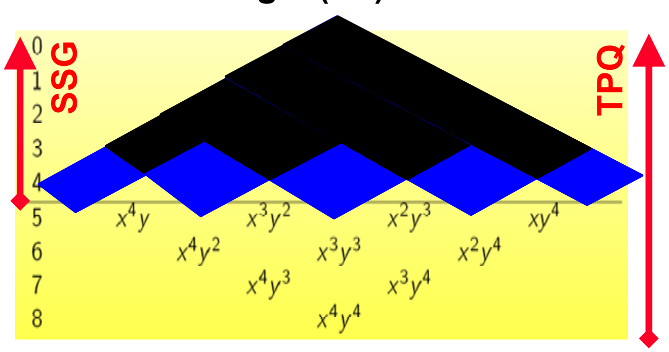
- **Every combination of 1-D rules**
- **Scales as m^n**
- **1-D Gaussian rule of order m → integrands to order $2m - 1$**
- **Assuming $\boldsymbol{R}\,\Psi_j$ of order 2p, select $m = p + 1$**

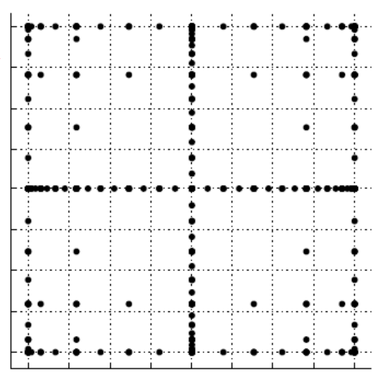


Smolyak Sparse Grid: PCE/SC

$$\boxed{\mathscr{A}(\mathrm{w},n) = \sum_{\mathrm{w}+1\le|\mathbf{i}|\le\mathrm{w}+n} (-1)^{\mathrm{w}+n-|\mathbf{i}|}\binom{n-1}{\mathrm{w}+n-|\mathbf{i}|}\cdot\left(\mathscr{U}^{i_1}\otimes\cdots\otimes\mathscr{U}^{i_n}\right)}$$

Pascal's triangle (2D):



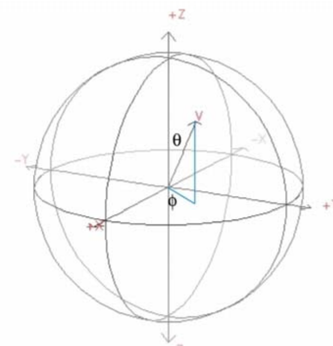


Cubature: PCE

Stroud and extensions (Xiu, Cools)

→ Low order PCE

→ global SA, anisotropy detection



Gaussian i = 2 → p = 1

$$\boxed{x_{k,2r-1} = \sqrt{2}\cos\frac{2rk\pi}{n+1}, \qquad x_{k,2r} = \sqrt{2}\sin\frac{2rk\pi}{n+1}}$$

Arbitrary PDF $\boxed{t^{(k)} = \frac{1}{\gamma}\left[\sqrt{\gamma c_1}x^{(k)} - \delta\right]}$

Sandia National Laboratories

Stochastic Expansion Example

- **Compare PCE to sampling and local reliability methods (modify `dakota_cantilever_lhs.in` or `dakota_cantilever_rel.in`)**

Example Input/Output: PCE

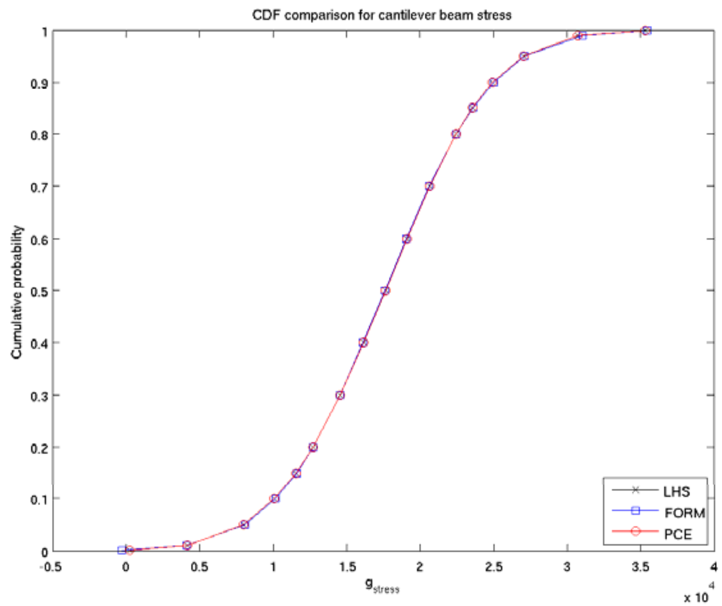
Input (extra_examples/dakota_uq_cantilever_pce.in)

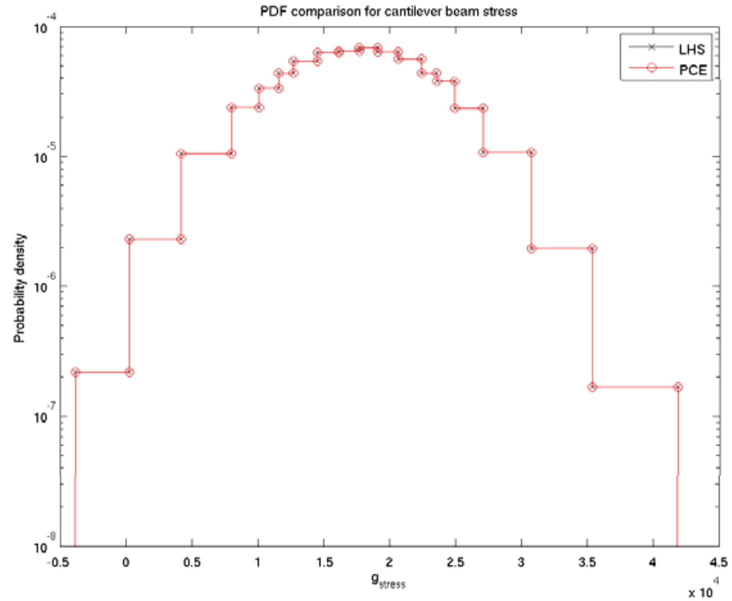
```
strategy
     single_method graphics

method,
     polynomial_chaos
       sparse_grid_level = 2 #non_nested
       sample_type lhs seed = 12347 samples = 10000
       num_probability_levels = 0 17 17
       probability_levels =
      .001 .01 .05 .1 .15 .2 .3 .4 .5 .6 .7 .8 .85 .9 .95 .99 .999
      .001 .01 .05 .1 .15 .2 .3 .4 .5 .6 .7 .8 .85 .9 .95 .99 .999
       cumulative distribution
#      output silent

variables,
     continuous_design = 2
       initial_point    2.5    2.5
       upper_bounds    10.0   10.0
       lower_bounds     1.0    1.0
       descriptors      'beam_width' 'beam_thickness'
     normal_uncertain = 4
       means           =  40000. 29.E+6 500. 1000.
       std_deviations =  2000. 1.45E+6 100. 100.
       descriptors       =  'R' 'E' 'X' 'Y'

interface,
     direct
       analysis_driver = 'mod_cantilever'

responses,
     descriptors = 'area' 'g_stress' 'g_displ'
     num_response_functions = 3
     no_gradients
     no_hessianse
```

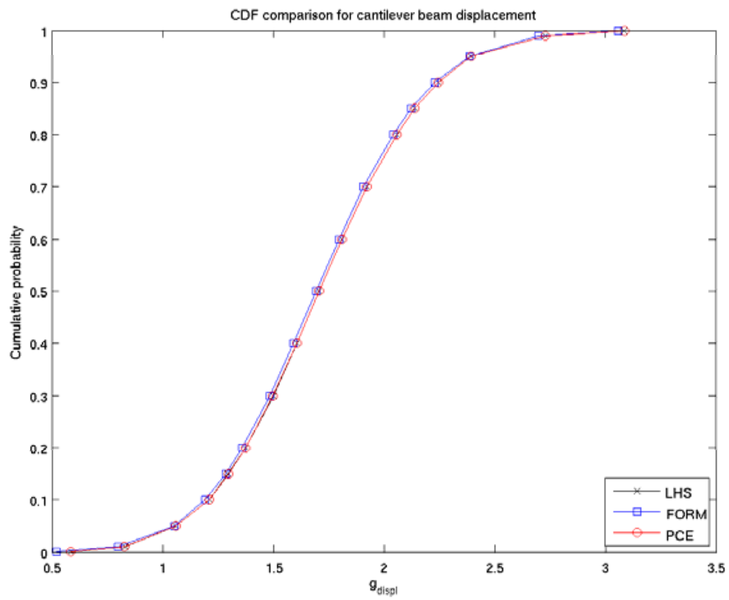
Output

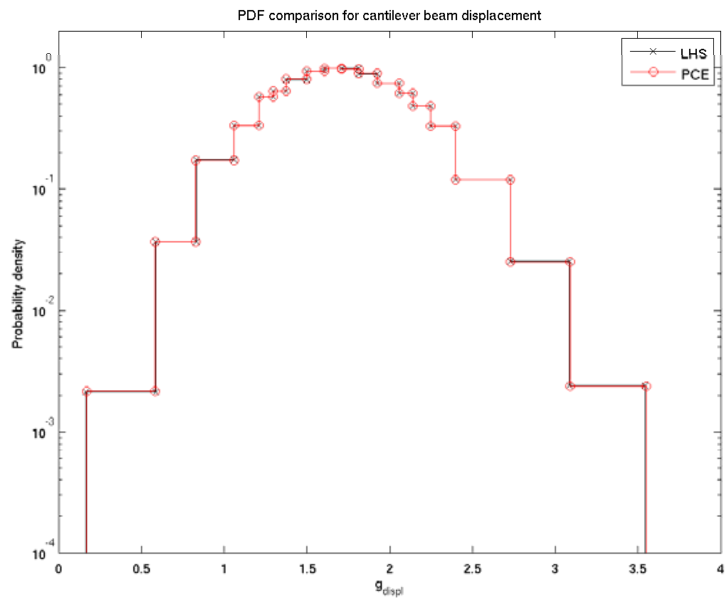
```
<<<<< Function evaluation summary: 57 total (57 new, 0 duplicate)
-------------------------------------------------------------------
Polynomial Chaos coefficients for area:
        coefficient   u1   u2   u3   u4
        ----------- ---- ---- ---- ----
   6.2500000000e+00  He0  He0  He0  He0   ...
Polynomial Chaos coefficients for g_stress:
        coefficient   u1   u2   u3   u4
        ----------- ---- ---- ---- ----
   1.7600000000e+04  He0  He0  He0  He0 ...
Polynomial Chaos coefficients for g_displ:
        coefficient   u1   u2   u3   u4
        ----------- ---- ---- ---- ----
   1.7201243431e+00  He0  He0  He0  He0 ...

Statistics derived analytically from polynomial expansion:

Moment-based statistics for each response function:
                            Mean           Std Dev          Skewness          Kurtosis
area
  expansion:    6.2500000000e+00  2.4824701829e-15
  numerical:    6.2500000000e+00  7.1054273576e-15  1.0000000000e+00 -2.0000000000e+00
g_stress
  expansion:    1.7600000000e+04  5.7871581973e+03
  numerical:    1.7600000000e+04  5.7871581973e+03  9.4100742175e-15  6.2172489379e-15
g_displ
  expansion:    1.7201243431e+00  4.0644795983e-01
  numerical:    1.7201243431e+00  4.0644787032e-01  1.5009952217e-01  4.9005496977e-02

Statistics based on 10000 samples performed on polynomial expansion:

Cumulative Distribution Function (CDF) for g_stress:
     Response Level  Probability Level  Reliability Index  General Rel Index
     --------------  -----------------  -----------------  -----------------
   2.4921421856e+02   1.0000000000e-03
   4.1489075797e+03   1.0000000000e-02
   ...
   3.0752664401e+04   9.9000000000e-01
   3.5331778223e+04   9.9900000000e-01
Cumulative Distribution Function (CDF) for g_displ:
     Response Level  Probability Level  Reliability Index  General Rel Index
     --------------  -----------------  -----------------  -----------------
   5.8392829293e-01   1.0000000000e-03
   8.2796204947e-01   1.0000000000e-02
   ...
   2.7290918315e+00   9.9000000000e-01
   3.0882954345e+00   9.9900000000e-01
```

Example Comparison







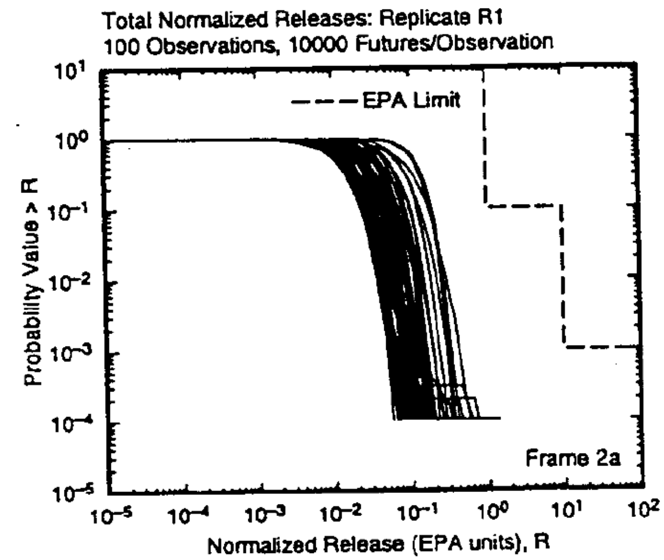


TIME PERMITTING

Epistemic UQ

Epistemic UQ: one does not know enough to specify probability distributions

Sometimes referred to as subjective, reducible, or lack of knowledge uncertainty

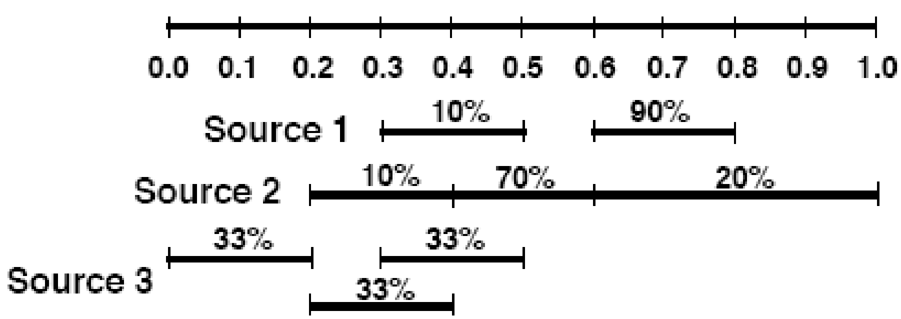


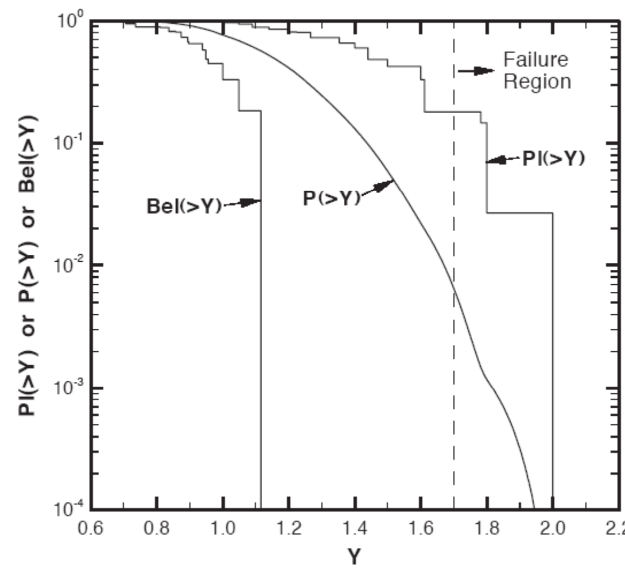
Second-order probability

– Two levels: distributions/intervals on distribution parameters
– Outer level can be epistemic (e.g., interval)
– Inner level can be aleatory (probability distrs)
– Strong regulatory history (NRC, WIPP).

Dempster-Shafer theory of evidence

– Basic probability assignment (interval-based)
– Solve opt. problems (currently sampling-based) to compute belief/plausibility for output intervals

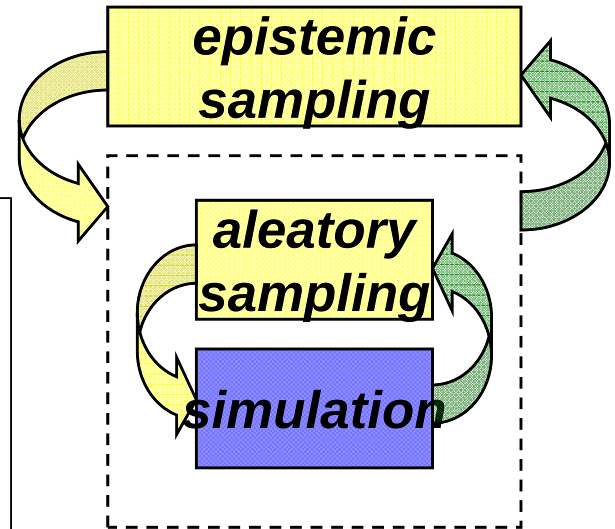




Imprecise probability (p-boxes), Info gap, …

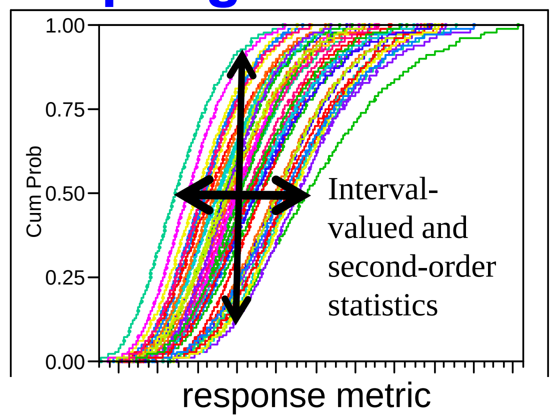
Mixed Aleatory-Epistemic UQ: IVP, DSTE, and SOP

Epistemic uncertainty (aka: subjective, reducible, lack of knowledge uncertainty): insufficient info to specify objective probability distributions



Traditional approach: nested sampling

- Expensive sims → under-resolved sampling (especially @ outer loop)
- Under-prediction of credible outcomes



Algorithmic approaches

- Interval-valued probability (IVP), *aka* PBA
- Dempster-Shafer theory of evidence (DSTE)
- Second-order probability (SOP), *aka* PoF

Increasing epistemic structure (stronger assumptions)

Address accuracy and efficiency

- Inner loop: stochastic exp. that are epistemic-aware (aleatory, combined)
- Outer loop:
  - IVP, DSTE: opt-based interval estimation, global (EGO) or local (NLP)
  - SOP: nested stochastic exp. (nested expectation is only post-processing in special cases)

$$\begin{array}{ll} \text{minimize} & M(s) \\ \text{subject to} & s_L \leq s \leq s_U \end{array}$$

$$\begin{array}{ll} \text{maximize} & M(s) \\ \text{subject to} & s_L \leq s \leq s_U \end{array}$$

Additional Resources

References

- Full list of research publications: http://dakota.sandia.gov/publications.html
- Selected UQ algorithm publications:
  - Local reliability: http://dakota.sandia.gov/papers/AIAA-2006-1828.pdf
  - Global reliability: http://dakota.sandia.gov/papers/AIAA-2007-1946.pdf
  - Stochastic expansions: http://dakota.sandia.gov/papers/AIAA_MAO2010_Adapt.pdf, recent *IJ4UQ* to appear
  - Epistemic: http://dakota.sandia.gov/papers/294_swi.pdf, recent *RESS* to appear
  - Application examples: http://dakota.sandia.gov/applications.html
- DAKOTA method documentation: http://dakota.sandia.gov/documentation.html (see Ch. 6 of Users Manual for UQ)

Software Downloads

- DAKOTA: http://dakota.sandia.gov/download.html
- Related packages: http://dakota.sandia.gov/packages.html

Sandia National Laboratories