Monte Carlo Simulation
with
Imprecise Random Variables

Thomas Fetz, Michael Oberguggenberger, Martin Schwarz

Unit for Engineering Mathematics
University of Innsbruck, Austria

{Thomas.Fetz, Michael.Oberguggenberger, Martin.Schwarz}@uibk.ac.at

IPW2015, Liverpool, UK

Problem statement and outline of presentation

Given

- Expensive input-output map $g : \mathbb{R}^n \to \mathbb{R} : x \to g(x)$.

  E.g. finite element computations (minutes or hours per computation).
- Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables modelling the uncertainty of variable x.

Aim

- Upper/lower probabilities that $g(x) \in B$.
- Upper/lower probabilities that $g(x) \leq y$. (upper/lower cumulative distribution function, p-box)
- Upper/lower probabilities that $g(x) \leq 0$. (upper/lower probability of failure)

Two approaches

- Monte-Carlo simulation of the family $\{g(X_\lambda)\}_{\lambda\in\Lambda}$.
- Monte-Carlo simulation of the random set $\mathcal{X}$ generated by $\{g(X_\lambda)\}_{\lambda\in\Lambda}$.

Numerical example

- The efficiency of the two approaches is demonstrated by means of a moderate scale engineering structure (simplified model of ARIANE 5 front skirt).

Two approaches

1 Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

- Probability space (Ω, Σ, m).
- Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

$$X_\lambda : \Omega \to \mathbb{R} : \omega \to X_\lambda(\omega).$$

- Probability $P(X_\lambda \in B)$ for fixed λ:

$$P(X_\lambda \in B) = \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B} \,\mathrm{d}m(\omega).$$

(for initial analysis we drop the map g)

2 Random set $\mathcal{X}$ based on $\{X_\lambda\}_{\lambda\in\Lambda}$

- Set-valued map $\mathcal{X} : \Omega \to \mathbb{R}$ defined by

$$\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\}.$$

(focal set at fixed ω)

- $\mathcal{X}$ is a random set, with upper/lower inverses

$$\mathcal{X}^-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \cap B \neq \varnothing\},$$
$$\mathcal{X}_-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \subseteq B\}.$$

Two approaches

1 Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

- Probability space (Ω,Σ,m).
- Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

$$X_\lambda : \Omega \to \mathbb{R} : \omega \to X_\lambda(\omega).$$

- Probability $P(X_\lambda \in B)$ for fixed λ:

$$P(X_\lambda \in B) = \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{d}m(\omega).$$

(for initial analysis we drop the map g)

2 Random set $\mathcal{X}$ based on $\{X_\lambda\}_{\lambda\in\Lambda}$

- Set-valued map $\mathcal{X} : \Omega \to \mathbb{R}$ defined by

$$\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\}.$$

(focal set at fixed ω)

- $\mathcal{X}$ is a random set, with upper/lower inverses

$$\mathcal{X}^-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \cap B \neq \varnothing\},$$
$$\mathcal{X}_-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \subseteq B\}.$$

Lower/upper probabilities for $\{X_\lambda\}_{\lambda\in\Lambda}$

$$\underline{P}(B) = \inf_{\lambda\in\Lambda} P(X_\lambda \in B) = \inf_{\lambda\in\Lambda} \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{d}m(\omega)$$

$$\overline{P}(B) = \sup_{\lambda\in\Lambda} P(X_\lambda \in B) = \sup_{\lambda\in\Lambda} \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{d}m(\omega)$$

Lower/upper probabilities for $\mathcal{X}$

$$\underset{\sim}{P}(B) = m(\mathcal{X}_-(B)) = \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\subseteq B}\,\mathrm{d}m(\omega)$$

$$\widetilde{P}(B) = m(\mathcal{X}^-(B)) = \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\cap B\neq\varnothing}\,\mathrm{d}m(\omega)$$

Two approaches

1 Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

- Probability space (Ω, Σ, m).
- Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

$$X_\lambda : \Omega \to \mathbb{R} : \omega \to X_\lambda(\omega).$$

- Probability $P(X_\lambda \in B)$ for fixed λ:

$$P(X_\lambda \in B) = \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\, \mathrm{d}m(\omega).$$

(for initial analysis we drop the map g)

2 Random set $\mathcal{X}$ based on $\{X_\lambda\}_{\lambda\in\Lambda}$

- Set-valued map $\mathcal{X} : \Omega \to \mathbb{R}$ defined by

$$\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\}.$$

(focal set at fixed ω)

- $\mathcal{X}$ is a random set, with upper/lower inverses

$$\mathcal{X}^-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \cap B \neq \varnothing\},$$
$$\mathcal{X}_-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \subseteq B\}.$$

Lower/upper probabilities for $\{X_\lambda\}_{\lambda\in\Lambda}$

$$\underline{P}(B) = \inf_{\lambda\in\Lambda} P(X_\lambda \in B) = \inf_{\lambda\in\Lambda} \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\, \mathrm{d}m(\omega)$$

$$\overline{P}(B) = \sup_{\lambda\in\Lambda} P(X_\lambda \in B) = \sup_{\lambda\in\Lambda} \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\, \mathrm{d}m(\omega)$$

Lower/upper probabilities for $\mathcal{X}$

$$\underset{\sim}{P}(B) = m(\mathcal{X}_-(B)) = \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\subseteq B}\, \mathrm{d}m(\omega)$$

$$\widetilde{P}(B) = m(\mathcal{X}^-(B)) = \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\cap B\neq\varnothing}\, \mathrm{d}m(\omega)$$

Theorem

$$\underset{\sim}{P} \le \underline{P} \le \overline{P} \le \widetilde{P}$$ $\mathcal{X}$ is more conservative than $\{X_\lambda\}_{\lambda\in\Lambda}$!

Simulation of a family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables

1 Basic sample $x_1,\dots,x_{N_{\text{samp}}}$

- Generate a sample $x_1,\dots,x_{N_{\text{samp}}}$ which is distributed as a **basic random variable** X_*.
- Distribution of X_* should cover a greater range than a distribution of a single X_λ does.

2 N_{samp} function evaluations $g(x_k)$, $k=1,\dots,N_{\text{samp}}$

- We compute $g(x_k)$ either using g directly or a cost saving surrogate model $\tilde{g}$.

3 Approximation of $P(g(X_\lambda)\le y)$

- Probability $P(g(X_\lambda)\le y)$ for fixed λ is computed by **reweighting** the original sample.
- Weights $w_k(\lambda)$ depending on parameters λ for reweighting the sample $x_1,\dots,x_{N_{\text{samp}}}$ according to the distribution of X_λ (cf. importance sampling):

$$w_k(\lambda)=\frac{f_{X_\lambda}(x_k)}{f_{X_*}(x_k)}\,\frac{1}{N_{\text{samp}}}=\frac{f_{\text{new}}(x_k)}{f_{\text{old}}(x_k)}\,\frac{1}{N_{\text{samp}}}.$$

- Obtaining $P(g(X_\lambda)\le y)$ for different X_λ **without additional function evaluations** of g:

$$P(g(X_\lambda)\le y)=\int_\Omega \mathbb{1}_{g(X_\lambda(\omega))\le y}\,\mathrm{d}m(\omega)\approx\sum_{k=1}^{N_{\text{samp}}}\mathbb{1}_{g(X_\lambda(\omega_k))\le y}\cdot w_k(\lambda)=\sum_{k=1}^{N_{\text{samp}}}\mathbb{1}_{g(x_k)\le y}\cdot w_k(\lambda).$$

4 Approximation of upper/lower cumulative distribution functions $\overline{F}(y)$ and $\underline{F}(y)$

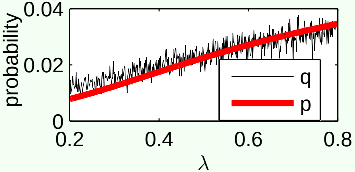
- We have to solve the following global optimization problems:

$$\overline{F}(y) = \overline{P}(g \le y) = \sup_{\lambda\in\Lambda} P(g(X_\lambda) \le y) \approx \max_{\lambda\in\Lambda} \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(x_k)\le y} \cdot w_k(\lambda) = \max_{\lambda\in\Lambda} p(\lambda),$$

$$\underline{F}(y) = \underline{P}(g \le y) = \inf_{\lambda\in\Lambda} P(g(X_\lambda) \le y) \approx \min_{\lambda\in\Lambda} \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(x_k)\le y} \cdot w_k(\lambda) = \min_{\lambda\in\Lambda} p(\lambda).$$

- **Smooth** and **cheap** objective function for fixed y: $p(\lambda) = \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(x_k)\le y} \cdot w_k(\lambda).$
- Standard optimization algorithms can be applied because of the smoothness of $p(\lambda)$.
- **Effort:** $N_{\text{obj}} \cdot N_{\text{samp}}$ reweightings, N_{samp} expensive function evaluations of g.
- Remark: Different samples $x_1(\lambda), \ldots, x_{N_{\text{samp}}}(\lambda)$ for different $\lambda \in \Lambda$ would lead to a non-smooth objective function

$$q(\lambda) = \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(x_k(\lambda))\le y} \cdot \frac{1}{N_{\text{samp}}}.$$



Simulation of a random set $\mathcal{X}$

1 Propagation of a random set through g

- $\mathcal{G}(\omega) = g(\mathcal{X}(\omega)) = \{g(X_\lambda(\omega))) : \lambda \in \Lambda\}$
- $\mathcal{G}(\omega) = [\underline{\mathcal{G}}(\omega), \overline{\mathcal{G}}(\omega)]$ random interval
- $\underline{\mathcal{G}}(\omega) = \min g(\mathcal{X}(\omega)),\ \overline{\mathcal{G}}(\omega) = \max g(\mathcal{X}(\omega))$

2 Cumulative distribution functions

- $\widetilde{F}(y) = \widetilde{P}(g \le y),\ \underset{\sim}{F}(y) = \underset{\sim}{P}(g \le y)$
- $\widetilde{F}(y) = P\big((-\infty, y] \cap [\underline{\mathcal{G}}, \overline{\mathcal{G}}] \neq \varnothing\big) = P\big(\underline{\mathcal{G}} \le y\big) = F_{\underline{\mathcal{G}}}(y)$
- $\underset{\sim}{F}(y) = P\big([\underline{\mathcal{G}}, \overline{\mathcal{G}}] \subset (-\infty, y]\big) = P\big(\overline{\mathcal{G}} \le y\big) = F_{\overline{\mathcal{G}}}(y)$

3 Algorithm for computing $\widetilde{F}(y)$

- Generate $\omega_1, \ldots, \omega_{N_{\text{samp}}}$ distributed as m.
- For each ω_k, estimate $\underline{\mathcal{G}}(\omega_k) \approx \min_i g(X_{\lambda_i}(\omega_k))$ using grid points $\lambda_1, \ldots, \lambda_{N_{\text{grid}}}$ on Λ.
- $\widetilde{F}(y) \approx \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{\underline{\mathcal{G}}(\omega_k) \le y} \cdot \frac{1}{N_{\text{samp}}}$.

Effort: $N_{\text{grid}} \cdot N_{\text{samp}}$ expensive evaluations of g.

4 Cost saving methods, approximation of g by **surrogate models** $\widetilde{g}_i$

Starting point: Collocation points x_j, $j = 1, \ldots, N_{\mathrm{coll}}$, in $\mathbb{R}^n$ and N_{coll} evaluations $y_j = g(x_j)$.

Stochastic surrogate models $\widetilde{g}_i$ of maps $\Omega \to g \circ X_{\lambda_i}$:

- Collocation points x_j are pulled back to probability space Ω, i.e., for each λ_i and x_j, a collocation point $\omega_{ij} = X_{\lambda_i}^{-1}(x_j)$ in Ω is computed.
- Clearly, $y_j = g(X_{\lambda_i}(\omega_{ij})) = g(x_j)$ for every i.
- Stochastic surrogate models $\widetilde{g}_i$, $i = 1, \ldots, N_{\text{grid}}$, are obtained by regression through the data points (ω_{ij}, y_j), $j = 1, \ldots, N_{\text{coll}}$.
- Finally, one computes $\underline{\mathcal{G}}(\omega) \approx \min_{i=1,\ldots,N_{\text{grid}}} \widetilde{g}_i(\omega)$.
- Based on a sample $\omega_1, \ldots, \omega_{N_{\text{coll}}}$, a Monte Carlo sample of $\underline{\mathcal{G}}$ is obtained.

Effort: N_{coll} expensive evaluations of g;
N_{grid} linear regressions (moderate cost);
N_{samp} cheap evaluations of $\widetilde{g}_i$ for each i.

Advantage of stochastic surrogate models $\widetilde{g}_i$ on Ω:

- Use of orthogonal polynomials with respect to the measure m.
- In the Gaussian case, Ω is standard Gaussian space of dimension $n \times$(polynomial order) and $X_{\lambda_i}^{-1}(x_j)$ is simply $(x_j - \mu_i)/\sigma_i$ (in each component of x_j).

Limit state function:

- $g(x) = 1 - \max\{\mathsf{PEEQ}(x)/0.07, \mathsf{SP}(x)/155\}$
- PEEQ: maximum value of equivalent plastic strain.
- SP: maximum principal stress.

Parameters:

- 35 parameters in total.
- We model the uncertainty of the 3 most significant parameters using families $\{X_{(\mu_i,\sigma_i)}\}$ of Gaussian random variables.

	description	μ_i	σ_i^2
x_1	yield stress in cylinder 3	$[350, 375]\,\mathrm{N/mm^2}$	$[0.01, 0.02]$
x_2	pressures loads in sphere 2	$[0.38, 0.41]\,\mathrm{N/mm^2}$	$[0.01, 0.02]$
x_3	temperature loads in cylinder 1	$[430, 470]\,\mathrm{K}$	$[0.01, 0.02]$

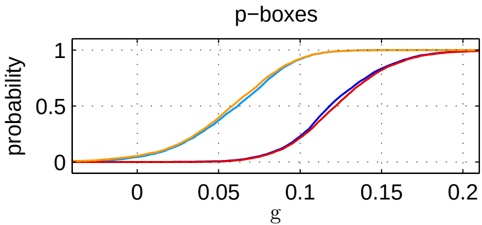
- The random variables are independent.
- Family $\{X_\lambda\}_{\lambda\in\Lambda}$ of joint random variables, $\lambda = (\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2)$

  $\Lambda = [350, 375] \times [0.38, 0.41] \times [430, 470] \times [0.01, 0.02] \times [0.01, 0.02] \times [0.01, 0.02]$.

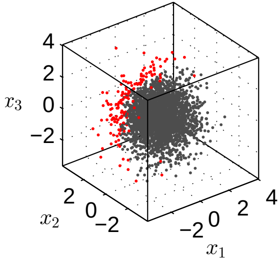
Numerical example: results

Upper probabilities: $\overline{p}_f = \overline{F}(0) = 0.0441,\ \widetilde{p}_f = \widetilde{F}(0) = 0.0554.$

Lower/upper cumulative distribution functions $\underset{\sim}{F} \le \underline{F} \le \overline{F} \le \widetilde{F}$**:**



Lower bounds $\underline{\mathcal{G}}$ **of random intervals** $\mathcal{G}$**:**



Red dots: $\omega_k : \underline{\mathcal{G}}(\omega_k) \le 0,$

Gray dots: $\omega_k : \underline{\mathcal{G}}(\omega_k) > 0.$

$$\widetilde{p}_f = \widetilde{F}(0) = \frac{\#\text{ red dots}}{N_{\text{samp}}}.$$

Conclusion

Two interpretations of imprecise probability models

(1) Information given by a family $\{g(X_\lambda)\}_{\lambda\in\Lambda}$.

(2) Information given by the random set $\mathcal{X}$ generated by the family.

Two numerical methods

(1) Reweighting a basic sample.

(2) Fitting stochastic surrogate models to all $g(X_\lambda)$.

Optimization

(1) and (2).

Numerical effort

(1) $<$ (2).

Choice of the method

Depends on interpretation of imprecise model.