JdS 2021: Active learning strategy for fragility curve estimation using adaptive importance sampling

C. Gauchy, C. Feau, J. Garnier

clement.gauchy@cea.fr; cyril.feau@cea.fr; josselin.garnier@polytechnique.edu

– June 7, 2021 –

Sommaire

Introduction

Seismic hazard

Izmir seism picture (from newspaper *Le Dauphiné*, November 1st 2020)

Seismic hazard

Izmir seism picture (from newspaper *Le Dauphiné*, November 1st 2020)

- Seisms have a huge social & human costs and are unfortunately random.
- Dramatic effect on mechanical structures.

Seismic hazard

Crucial need of seismic risk assesment for valuable industrial assets (e.g. nuclear power plant)

Nuclear power plant of Nogent sur Seine, France

Seismic probabilistic risk assesment

A cost-benefit analysis between safety of crucial mechanical structures (e.g. nuclear power plant, dam) and uncertainties tainting seismic ground motion. It can be sumed up in 3 main steps [1]:

- 1) Estimation of the probability density function $p(a)$ of the intensity measure of a seism a.
- 2) (my job) Estimation of the conditional probability of failure given the seismic intensity a: the **fragility curve** $P_f(a)$
- 3) Estimation of a probability of failure:

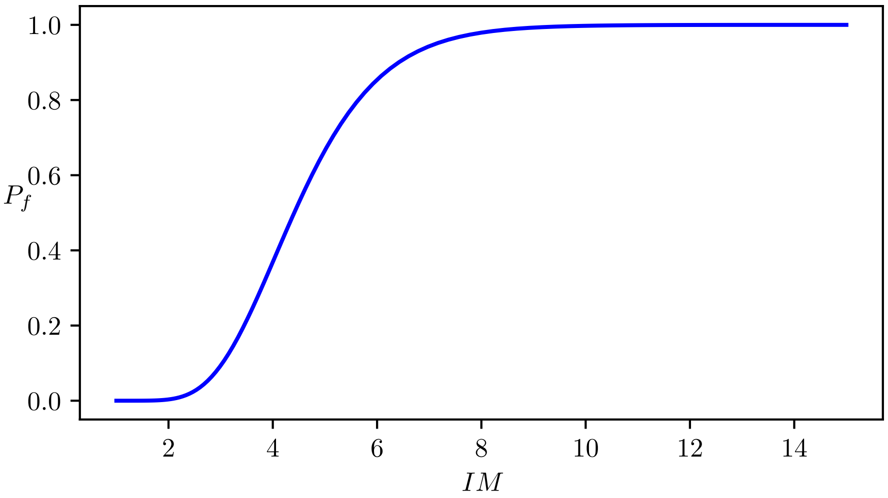
$$\psi = \int P_f(a)p(a)da$$

[1] Irmela Zentner, Max Gündel, and Nicolas Bonfils. Fragility analysis methods: Review of existing approaches and application. *Nuclear Engineering and Design*, 323:245 – 258, 2017.
ISSN 0029-5493.
doi: https://doi.org/10.1016/j.nucengdes.2016.12.021.
URL http://www.sciencedirect.com/science/article/pii/S0029549316305209

Seismic fragility curve

$$P_f(a) = \mathbb{P}(L > C | IM = a)$$

- L: mechanical demand of the structure, obtained using time-consuming numerical simulations.
- C critical level where the structure is considered in a failure state.
- IM *Intensity Measure* of a seismic ground motions.



Statistical learning of fragility curve

- Generate the data (X_i, S_i) where $S_i = \mathbb{1}_{L_i \geq C}$ and $X_i = \log(IM_i)$, using a computer model as an oracle. We use a synthetic generator of seismic signals to obtain X_i (e.g. a filtered modulated white noise).

- We suppose a lognormal shape for the fragility curve. This is a common parametric model used by mechanical engineers in seismic risk analysis:

$$\mathbb{P}(L > C | IM = a) = f_{\alpha,\beta}(a) = \Phi\left(\frac{\log\left(\frac{a}{\alpha}\right)}{\beta}\right)$$

- The parameters α and β are estimated by empirical risk minimization using least squares:

$$\widehat{\alpha}_n, \widehat{\beta}_n = \operatorname{argmin} \sum_{i=1}^{n} (S_i - f_{\alpha,\beta}(X_i))^2 \tag{1}$$

Adaptive importance sampling

Motivation

- The labels $S_i = \mathbb{1}_{L_i \geq C}$ are gathered using time-consuming mechanical simulation (e.g. Finite Elements Method).

[2] Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010.
doi: 10.1002/eqe.997.
URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.997`

Motivation

- The labels $S_i = \mathbb{1}_{L_i \geq C}$ are gathered using time-consuming mechanical simulation (e.g. Finite Elements Method).
- The log seismic intensity X_i are derived from synthetic ground motions [2] which are considered cheap to generate.

[2] Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010.
doi: 10.1002/eqe.997.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.997

Motivation

- The labels $S_i = \mathbb{1}_{L_i \geq C}$ are gathered using time-consuming mechanical simulation (e.g. Finite Elements Method).
- The log seismic intensity X_i are derived from synthetic ground motions [2] which are considered cheap to generate.
- **Goal: Minimize the sample size and maximize the reduction of uncertainties concerning the fragility curve**. → Active Learning

[2]Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010.
doi: 10.1002/eqe.997.
URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.997`

Principle

- Denote $\theta = (\alpha, \beta)^T$

Principle

- Denote $\theta = (\alpha, \beta)^T$
- Empirical risk minimization:

$$\widehat{R}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}(S_i - f_\theta(X_i))^2$$

Principle

- Denote $\theta = (\alpha, \beta)^T$
- Empirical risk minimization:

$$\widehat{R}_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} (S_i - f_\theta(X_i))^2$$

- Assume $X \sim p$ where p is known.

Principle

- Denote $\theta = (\alpha, \beta)^T$
- Empirical risk minimization:

$$\widehat{R}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}(S_i - f_\theta(X_i))^2$$

- Assume $X \sim p$ where p is known.
- **Idea:** Sample $X \sim q$ so that it minimizes $\mathrm{Var}(\widehat{R}_n(\theta)) \to$ Importance Sampling

Principle

- Denote $\theta = (\alpha, \beta)^T$
- Empirical risk minimization:

$$\widehat{R}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}(S_i - f_\theta(X_i))^2$$

- Assume $X \sim p$ where p is known.
- **Idea:** Sample $X \sim q$ so that it minimizes $\mathrm{Var}(\widehat{R}_n(\theta)) \to$ Importance Sampling
- IS-Empirical risk minimization:

$$\tilde{R}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}\frac{p(X_i)}{q(X_i)}(S_i - f_\theta(X_i))^2 \ .$$

Optimal sampling density

- $$\mathbb{V}(\tilde{R}_n(\theta))) = \frac{1}{n}\int\int \frac{p(x)^2}{q(x)}(s - f_\theta(x))^4 p(ds|x)dx - \mathbb{E}[(S - f_\theta(X))^2]^2$$

Optimal sampling density

- $$\mathbb{V}(\tilde{R}_n(\theta))) = \frac{1}{n}\int\int \frac{p(x)^2}{q(x)}(s - f_\theta(x))^4 p(ds|x)dx - \mathbb{E}[(S - f_\theta(X))^2]^2$$

- $$q_\theta^* = \underset{q}{\operatorname{argmin}}\, \mathbb{V}(\tilde{R}_n(\theta)) \ .$$

Optimal sampling density

- $$\mathbb{V}(\tilde{R}_n(\theta)) = \frac{1}{n}\int\int \frac{p(x)^2}{q(x)}(s - f_\theta(x))^4 p(ds|x)dx - \mathbb{E}[(S - f_\theta(X))^2]^2$$

- $$q_\theta^* = \underset{q}{\operatorname{argmin}}\, \mathbb{V}(\tilde{R}_n(\theta)) \ .$$

- $$q_\theta^* = \underset{q}{\operatorname{argmin}} \int_{x\in\mathcal{X}} \frac{p(x)^2}{q(x)}\tilde{\ell}_\theta(x)dx \ ,$$

- $\tilde{\ell}_\theta(x) = \mathbb{E}[(S - f_\theta(X))^4|X = x]$

$$q_\theta^*(x) \propto \sqrt{\tilde{\ell}_\theta(x)}p(x) \ ,$$

The optimal sampling density depends on θ ! $\rightarrow$ *Adaptive* importance sampling

AIS algorithm

Because P_f is unknown we perform the following approximation of $\tilde{\ell}_\theta(x)$:

$$\tilde{\ell}_\theta(x) \approx f_\theta(x)(1 - f_\theta(x))^4 + (1 - f_\theta(x))f_\theta(x)^4$$

AIS algorithm

Because P_f is unknown we perform the following approximation of $\tilde{\ell}_\theta(x)$:

$$\tilde{\ell}_\theta(x) \approx f_\theta(x)(1 - f_\theta(x))^4 + (1 - f_\theta(x))f_\theta(x)^4$$

The instrumental density chosen is:

$$q_\theta(x) \propto \sqrt{f_\theta(x)(1 - f_\theta(x))^4 + (1 - f_\theta(x))f_\theta(x)^4}\, p(x) \ ,$$

AIS algorithm

Because P_f is unknown we perform the following approximation of $\tilde{\ell}_\theta(x)$:

$$\tilde{\ell}_\theta(x) \approx f_\theta(x)(1 - f_\theta(x))^4 + (1 - f_\theta(x))f_\theta(x)^4$$

The instrumental density chosen is:

$$q_\theta(x) \propto \sqrt{f_\theta(x)(1 - f_\theta(x))^4 + (1 - f_\theta(x))f_\theta(x)^4}\, p(x) \ ,$$

Algorithm 3 Adaptive Importance Sampling

1. Initialisation: $\tilde{\theta}_0$
2. for $i = 1, \ldots, n$:
   2.1 Sample X_i from $q_{\tilde{\theta}_{i-1}}$
   2.2 Call the mechanical simulation at point X_i to get label S_i
   2.3 Compute

$$\tilde{\theta}_i = \underset{\theta}{\operatorname{argmin}} \frac{1}{i} \sum_{j=1}^{i} \frac{p(X_i)}{q_{\tilde{\theta}_{j-1}}(X_j)} (S_j - f_\theta(X_j))^2$$

AIS algorithm: numerical issues

- Actually not a good idea to use the sampling density $q^*_\theta(x)$ directly (uncontrolled likelihood ratio).

[3] Art Owen and Yi Zhou. Safe and effective importance sampling. *Journal of the American Statistical Association*, 95(449):135–143, 2000. ISSN 01621459. URL http://www.jstor.org/stable/2669533

AIS algorithm: numerical issues

- Actually not a good idea to use the sampling density $q^*_\theta(x)$ directly (uncontrolled likelihood ratio).
- *Defensive* AIS [3]:

$$q_{\theta,\varepsilon}(x) = \varepsilon p(x) + (1-\varepsilon) q^*_\theta(x)$$

- *Penalized* least squares for stability and identifiability issues:

$$\Omega(\theta; \beta_{reg}) = \frac{\beta_{reg}}{\beta}$$

$$\tilde{R}_{n,reg}(\theta; \beta_{reg}) = \frac{1}{n} \sum_{i=1}^{n} \frac{p(X_i)}{q_{\tilde{\theta}_{i-1},\varepsilon}(X_i)} (S_i - f_\theta(X_i))^2 + \frac{\Omega(\theta; \beta_{reg})}{n}$$

$$\tilde{\theta}_{n,\varepsilon} = \operatorname{argmin} \tilde{R}_{n,reg}(\theta; \beta_{reg})$$

[3] Art Owen and Yi Zhou. Safe and effective importance sampling. *Journal of the American Statistical Association*, 95(449):135–143, 2000. ISSN 01621459. URL http://www.jstor.org/stable/2669533

Theoretical results

Asymptotic normality of $\tilde{\theta}_n$

Theorem (Asymptotic normality of $\tilde{\theta}_n$)

Denote $\theta_* = \operatorname{argmin} \mathbb{E}[(S - f_\theta(X))^2]$. *Suppose* $\tilde{\theta}_n$ *consistent. With appropriate regularity conditions,* $\sqrt{n}(\tilde{\theta}_n - \theta_*)$ *is asymptotically normal with mean zero and covariance matrix* $\ddot{R}(\theta_*)^{-1} V(q_{\theta_*,\varepsilon})(\ddot{R}(\theta_*)^{-1})^T$ *with*

$$V(q_{\theta_*,\varepsilon}) = \mathbb{E}\left[\frac{p(X)}{q_{\theta_*,\varepsilon}(X)}(S - f_{\theta_*}(X))^2 \nabla f_{\theta_*}(X) \nabla f_{\theta_*}(X)^T\right]$$

Proof sketch

- Tools: M-estimation theory [4] and convergence of martingale [5].

[4] A. W. van der Vaart. *Asymptotic Statistics.*
Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.
doi: 10.1017/CBO9780511802256

[5] P. Hall, C.C. Heyde, Z.W. Birnbaum, and E. Lukacs. *Martingale Limit Theory and Its Application.*
Communication and Behavior. Elsevier Science, 2014.
ISBN 9781483263229.
URL `https://books.google.fr/books?id=gqriBQAAQBAJ`

Validity of the theorem

- For the lognormal parametric model $f_{\alpha,\beta}(a) = \Phi\left(\frac{\log\left(\frac{a}{\alpha}\right)}{\beta}\right)$. The theorem is valid as long as $(\alpha, \beta) \in \Theta$ where Θ is a compact in $\mathbb{R}^2$.
- Usually $\Theta = [\alpha_{min}, \alpha_{max}] \times [\beta_{min}, \beta_{max}]$

Stopping criterion

- Question: at which sample size does the parameter $\tilde{\theta}_n$ reaches asymptotic normality ?

Stopping criterion

- Question: at which sample size does the parameter $\tilde{\theta}_n$ reaches asymptotic normality ?
- Consider two independent datasets $\mathcal{D}_1 = (X_{i,1}, S_{i,1})_{i\in\{1,\ldots,n\}}$ and $\mathcal{D}_2 = (X_{i,2}, S_{i,2})_{i\in\{1,\ldots,n\}}$.

$$\tilde{\theta}_{n,j} = \operatorname*{argmin}_{\theta\in\Theta} \tilde{R}_{n,j}(\theta), \qquad j = 1, 2\,.$$

Stopping criterion

- Question: at which sample size does the parameter $\tilde{\theta}_n$ reaches asymptotic normality ?
- Consider two independent datasets $\mathcal{D}_1 = (X_{i,1}, S_{i,1})_{i\in\{1,\dots,n\}}$ and $\mathcal{D}_2 = (X_{i,2}, S_{i,2})_{i\in\{1,\dots,n\}}$.

$$\tilde{\theta}_{n,j} = \underset{\theta\in\Theta}{\operatorname{argmin}}\, \tilde{R}_{n,j}(\theta), \qquad j = 1, 2\,.$$

$$\sqrt{n}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1})) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 8V(q_{\theta_*,\varepsilon}))$$

Stopping criterion

- Question: at which sample size does the parameter $\tilde{\theta}_n$ reaches asymptotic normality ?
- Consider two independent datasets $\mathcal{D}_1 = (X_{i,1}, S_{i,1})_{i\in\{1,\dots,n\}}$ and $\mathcal{D}_2 = (X_{i,2}, S_{i,2})_{i\in\{1,\dots,n\}}$.

$$\tilde{\theta}_{n,j} = \operatorname*{argmin}_{\theta\in\Theta} \tilde{R}_{n,j}(\theta), \qquad j = 1, 2\,.$$

$$\sqrt{n}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1})) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 8V(q_{\theta_*,\varepsilon}))$$

- Stopping criterion:

$$\widehat{W}_n = \frac{n}{8}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1}))^T \widehat{V}_{n,12}^{-1}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1}))$$

$$\widehat{V}_{n,12} = \frac{1}{2}\left(\widehat{V}_{n,1}(\tilde{\theta}_{n,1}) + \widehat{V}_{n,2}(\tilde{\theta}_{n,2})\right)$$

$$\widehat{V}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \frac{p(X_i)^2}{q_{\theta,\varepsilon}(X_i) q_{\tilde{\theta}_{i-1},\varepsilon}(X_i)}(S_i - f_\theta(X_i))^2 \nabla f_\theta(X_i) \nabla f_\theta(X_i)^T$$

Stopping criterion

- Question: at which sample size does the parameter $\tilde{\theta}_n$ reaches asymptotic normality ?
- Consider two independent datasets $\mathcal{D}_1 = (X_{i,1}, S_{i,1})_{i\in\{1,\dots,n\}}$ and $\mathcal{D}_2 = (X_{i,2}, S_{i,2})_{i\in\{1,\dots,n\}}$.

$$\tilde{\theta}_{n,j} = \operatorname*{argmin}_{\theta\in\Theta} \tilde{R}_{n,j}(\theta), \qquad j = 1, 2\,.$$

$$\sqrt{n}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1})) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 8V(q_{\theta_*,\varepsilon}))$$

- Stopping criterion:

$$\widehat{W}_n = \frac{n}{8}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1}))^T \widehat{V}_{n,12}^{-1}(\dot{\tilde{R}}_{n,1}(\tilde{\theta}_{n,2}) - \dot{\tilde{R}}_{n,2}(\tilde{\theta}_{n,1}))$$

$$\widehat{V}_{n,12} = \frac{1}{2}\left(\widehat{V}_{n,1}(\tilde{\theta}_{n,1}) + \widehat{V}_{n,2}(\tilde{\theta}_{n,2})\right)$$

$$\widehat{V}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \frac{p(X_i)^2}{q_{\theta,\varepsilon}(X_i) q_{\tilde{\theta}_{i-1},\varepsilon}(X_i)}(S_i - f_\theta(X_i))^2 \nabla f_\theta(X_i) \nabla f_\theta(X_i)^T$$

- By Slutsky's lemma, $\widehat{W}_n$ converges weakly to $\chi^2(2)$.

Numerical results

Nonlinear oscillator

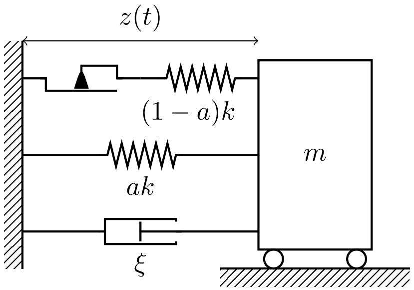
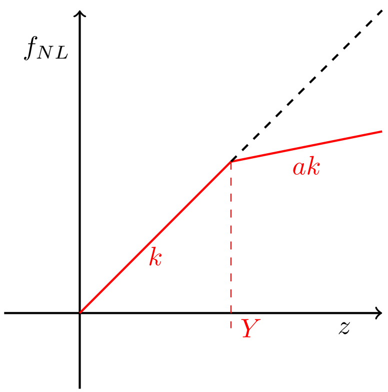


Figure: Elasto-plastic mechanical oscillator with kinematic hardening with parameter, $f_L = 5$ Hz, $\beta = 2\%$, the yield limit is $Y = 5.10^{-3}$ m and the post-yield stiffness is 20% of the elastic stiffness hence $a = 0.2$



$$\ddot{z}(t) + 2\beta\omega_L\dot{z}(t) + f_{NL}(t) = -s(t) \ ,$$

Nonlinear oscillator

- Consider a dataset of 10^5 sythetic seismic signal $t \to s_i(t)$.
- $z_i(t)$ displacement of the oscillator corresponding to signal s_i.

- A natural mechanical demand:
$$L_i = \max_{t \in [0,T]} |z_i(t)| \ .$$

Nonlinear oscillator

- Consider a dataset of 10^5 sythetic seismic signal $t \to s_i(t)$.
- $z_i(t)$ displacement of the oscillator corresponding to signal s_i.
- A natural mechanical demand:

$$L_i = \max_{t\in[0,T]} |z_i(t)| \ .$$

- Failure state for this oscillator:

$$S_i = \mathbb{1}_{(L_i > C)}$$

- $C = 2Y$ approximatly a 90% quantile of the mechanical response L.
- Our seismic intensity measure will be the peak ground acceleration (PGA):

$$PGA_i = \max_{t\in[0,T]} |s_i(t)|$$

Nonlinear oscillator: AIS benchmark setup

- Initialisation using a linear model approximation of the oscillator.

Nonlinear oscillator: AIS benchmark setup

- Initialisation using a linear model approximation of the oscillator.
- 20 points sampled using the instrumental density at the initial parameter θ_0.

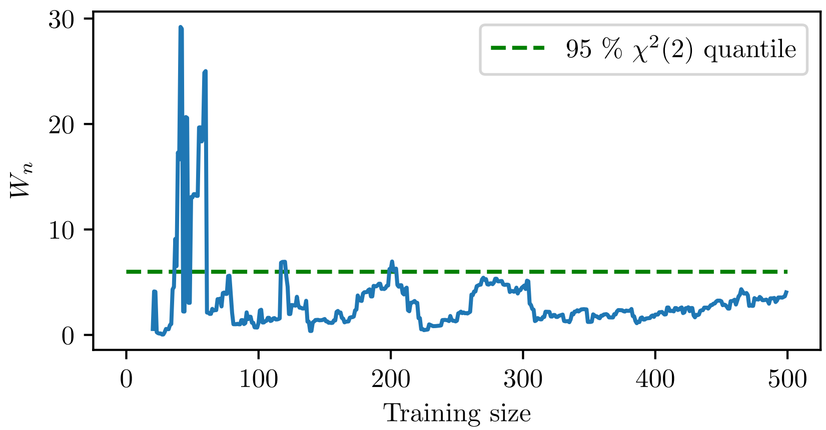
Nonlinear oscillator: AIS benchmark setup

- Initialisation using a linear model approximation of the oscillator.
- 20 points sampled using the instrumental density at the initial parameter θ_0.
- Penalization parameter β_{reg} chosen by cross-validation on the initial dataset of 20 datapoints.

Nonlinear oscillator: AIS benchmark setup

- Initialisation using a linear model approximation of the oscillator.
- 20 points sampled using the instrumental density at the initial parameter θ_0.
- Penalization parameter β_{reg} chosen by cross-validation on the initial dataset of 20 datapoints.
- 500 replications of 100 queries in the unlabeled pool of 10^5 signals using acceptation reject algorithm.

Stopping criterion for the nonlinear oscillator

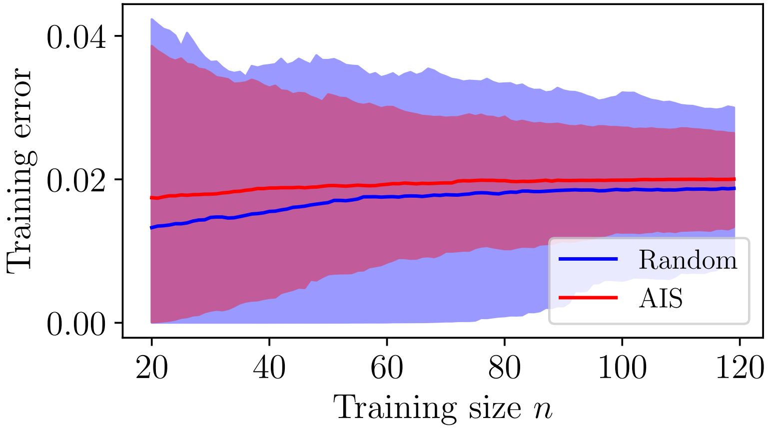


Value of the stopping criterion for two independent samples with varying sizes form 20 to 500.

The observed values of $\widehat{W}_n$ suggests to stop at $n = 100$ and then to build asymptotic confidence interval.

Nonlinear oscillator: numerical benchmark

Comparison betweeen AIS and sampling at random.



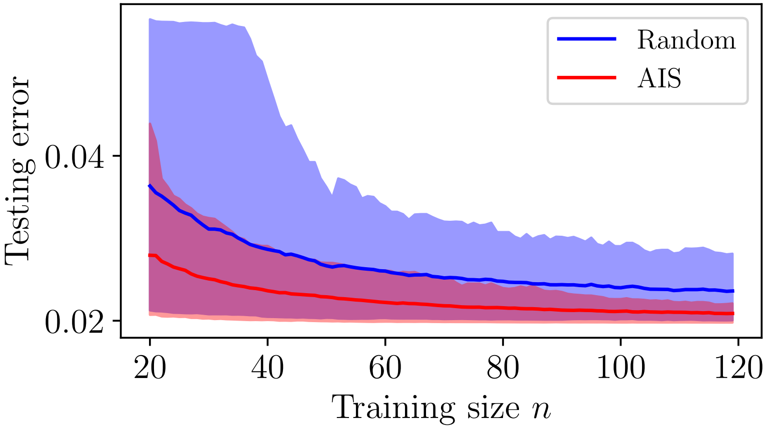
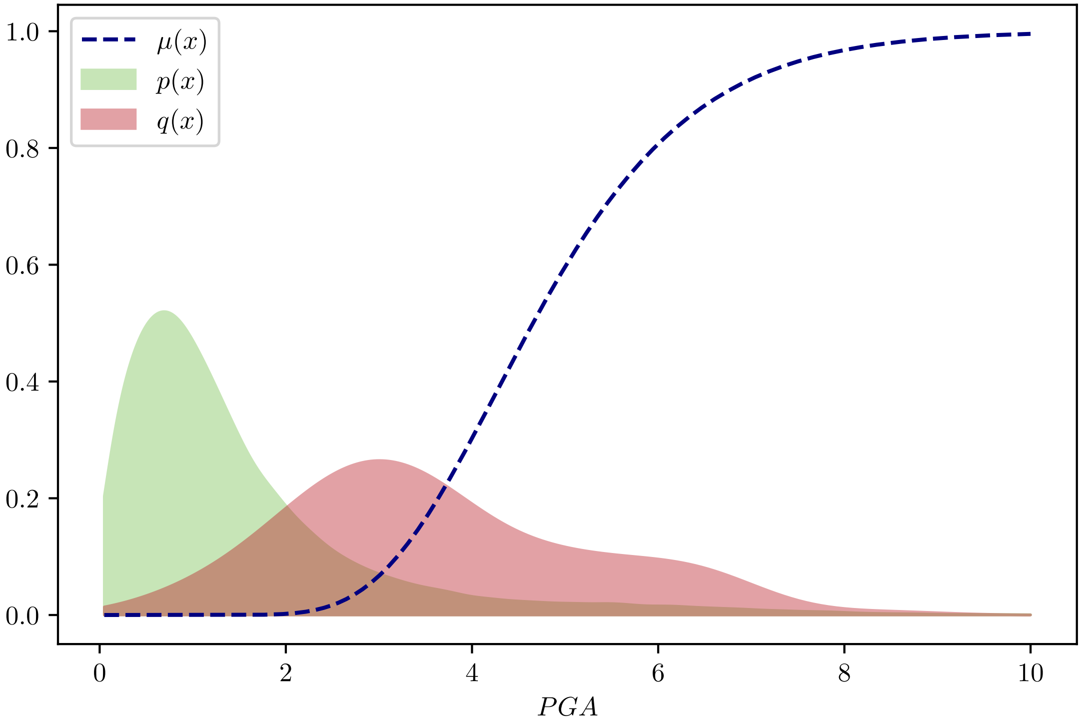


Figure: Numerical benchmark of the nonlinear oscillator.

Nonlinear oscillator: numerical benchmark



Estimation of the marginal density of PGA compared to the estimation of the density of sampled PGA using AIS with 100 datapoints for the nonlinear oscillator.

Nonlinear oscillator: numerical benchmark

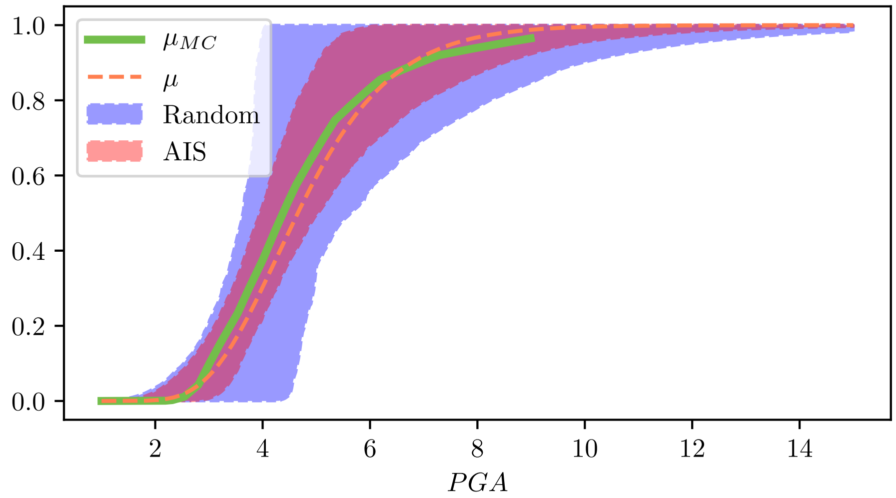


Figure: Estimation of the confidence intervals for random sampling and ADIS for the nonlinear oscillator

Conclusion & takeaways

- Adaptive importance sampling provide uncertainty reduction for estimation of fragility curves.

Conclusion & takeaways

- Adaptive importance sampling provide uncertainty reduction for estimation of fragility curves.
- Paper soon to be submitted in *Reliability assessment and safety engineering*.

Conclusion & takeaways

- Adaptive importance sampling provide uncertainty reduction for estimation of fragility curves.
- Paper soon to be submitted in *Reliability assessment and safety engineering*.
- **Program for the rest of the thesis**: Uncertainty propagation of the mechanical parameters of the structure, construction of a multi-fidelity metamodel.

Thank you for your attention !

Références

P. Hall, C.C. Heyde, Z.W. Birnbaum, and E. Lukacs. *Martingale Limit Theory and Its Application*. Communication and Behavior. Elsevier Science, 2014. ISBN 9781483263229. URL https://books.google.fr/books?id=gqriBQAAQBAJ.

Art Owen and Yi Zhou. Safe and effective importance sampling. *Journal of the American Statistical Association*, 95(449):135–143, 2000. ISSN 01621459. URL http://www.jstor.org/stable/2669533.

Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010. doi: 10.1002/eqe.997. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.997.

A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998. doi: 10.1017/CBO9780511802256.

Irmela Zentner, Max Gündel, and Nicolas Bonfils. Fragility analysis methods: Review of existing approaches and application. *Nuclear Engineering and Design*, 323:245 – 258, 2017. ISSN 0029-5493. doi: https://doi.org/10.1016/j.nucengdes.2016.12.021. URL http://www.sciencedirect.com/science/article/pii/S0029549316305209.

Merci pour votre attention !

Commissariat à l'énergie atomique et aux énergies alternatives
Centre de Saclay | 91191 Gif-sur-Yvette Cedex
T. +33 (0)1 XX XX XX XX

Direction de l'Energie Nucléaire
Département de Modélisation des Systèmes et Structures
Service d'Études Mécanique et Thermiques

Backup