Targeted Random Sampling for Reliability Assessment: A Demonstration of Concept

Michael D. Shields
Assistant Professor
Dept. of Civil Engineering
Johns Hopkins University

V.S. Sundar
Postdoctoral Fellow
Dept. of Civil Engineering
Johns Hopkins University
Motivation of Present Work

• Monte Carlo Simulation (MCS) is the most robust and accurate (and sometimes the only) means of reliability analysis
• MCS is often (rightfully so) criticized for its high computational cost
• The cost of MCS has been greatly reduced through the advent of sophisticated sampling and simulation methods
• We can still do better!
Outline of Presentation

• Overview of Reliability Methods
  – Emphasis on Monte Carlo Simulation based methods

• Refined Stratified Sampling
  – Concept / Theory
  – Algorithms / Applications

• Targeted Random Sampling
  – Concept / Theory
  – Algorithm
  – Examples

• Challenges & Future Directions
Reliability Analysis

Reliability Methods

Approximate Methods

- FORM
  - Low accuracy
  - Very efficient
  - Extensible to multiple failure modes
  - Moderate dimensions

- SORM
  - Moderate accuracy
  - Low efficiency
  - Single failure mode only
  - Moderate dimensions

- Response Surface
  - Moderate accuracy
  - Moderate efficiency
  - Multiple failure modes
  - Moderate to high dimensions

Exact Methods

- Monte Carlo
  - Very Robust
  - Accurate
  - Numerous methods available
  - Computationally expensive

- Numerical Int.
  - Computationally intractable for many cases
  - Only applicable in low dimension with special limit states.

- Direct MC
  - Robust
  - Accurate
  - Very Expensive
  - High dimensions
  - Multiple failure modes

- Importance Sampling
  - Moderate accuracy
  - Moderate efficiency
  - Widely used
  - Many variants

- Subset Simulation
  - Accurate
  - Efficient
  - High dimensions
  - MCMC can cause problems

- Line Sampling
  - Very accurate & efficient
  - High dimensions
  - Multiple failure modes

- Targeted Sampling

Stratified Sampling

- Sample size extension is straightforward
- No control over where new samples are placed
- No need to compute sample weights

Random Sampling

- Almost limitless control over how samples are added
- Carefully select, based on the available information, which strata to add samples to
- Is there a better way?
Stratified Sampling

- Sample size extension is straightforward
- No control over where new samples are placed
- No need to compute sample weights

Random Sampling

- Almost limitless control over how samples are added
- Carefully select, based on the available information, which strata to add samples to
- Is there a better way?

Stratified Sampling
Stratified Sampling

- Sample size extension is straightforward
- No control over where new samples are placed
- No need to compute sample weights

Random Sampling

- Almost limitless control over how samples are added
- Carefully select, based on the available information, which strata to add samples to
- Is there a better way?
Refined Stratified Sampling

1. Randomly sample

2. Randomly divide probability space into strata \( \Omega_i, i=1,2 \)

3. Randomly sample in empty stratum \( \Omega_i \) and assign weights: \( w_i = P(\Omega_i) \)

4. Randomly divide stratum \( \Omega_k \) to obtain strata \( \Omega_i; i=1,2,3 \)

5. Randomly sample in empty stratum \( \Omega_i \) and assign new weights: \( w_i = P(\Omega_i) \)

6. Divide largest stratum \( \Omega_k \) to obtain strata \( \Omega_i; i=1,2,3,4 \)

7. Randomly sample in empty stratum \( \Omega_i \) and assign new weights: \( w_i = P(\Omega_i) \)

8. Randomly divide stratum \( \Omega_i \) to obtain strata \( \Omega_i; i=1,2,3,4,5 \)

9. Randomly sample in empty stratum \( \Omega_i \) and assign new weights: \( w_i = P(\Omega_i) \)

Sample carried from previous step.
New stratum.
New sample.

\( X \sim \text{Probability space for RV1} \)
\( Y \sim \text{Probability space for RV2} \)

Why Refine Strata?

- Consider a \( N \) samples from stratum \( \Omega_1 \):

\[
Var\left[ T_S^1 \right] = \frac{p_1^2}{N} \sigma_1^2 + \sum_{j=1}^{M} \frac{p_j^2}{n_j} \sigma_j^2
\]

Other strata

\( \Omega_1 \)

- Consider stratum refinement with a single sample in each stratum:

\[
Var\left[ T_S^2 \right] = \frac{p_1^2}{N^2} \sum_{i=1}^{N} \sigma_i^2 + \sum_{j=1}^{M} \frac{p_j^2}{n_j} \sigma_j^2
\]

Other strata

\( \Omega_1 \)

- Evaluate the difference in the variances:

\[
Var\left[ T_S^1 \right] - Var\left[ T_S^2 \right] = \frac{p_1^2}{N} \left( \sigma_1^2 - \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right) > 0
\]

(Assumes a “balanced” sub-stratification)
Variance Savings by RSS: Example

• Consider samples drawn from $x \sim N(0,1)$ with $T_s = E[x]$:

\[
\begin{align*}
\text{Var}[T_s^b] &= \frac{1}{2} \sigma_1^2 + \frac{1}{2} \sigma_2^2 \\
\text{Var}[T_s^b] &= 0.090875 \\
\text{Var}[T_s^u] &= \frac{1}{4} \sigma_1^2 + \frac{1}{4} \sigma_2^2 + \frac{1}{4} \sigma_3^2 \\
\text{Var}[T_s^u] &= 0.0629 \\
\text{Var}[T_s^u] &= 0.0349 \\
\end{align*}
\]

$>30\%$ reduction in variance by adding one stratum

$>60\%$ reduction in variance by fully stratifying
RSS Algorithms

1. Random Stratum Division
2. Variance Minimization
3. Sensitivity-Based Stratum Division
4. Enhanced Space-Filling
5. Targeted Random Sampling
6. Others (Stochastic Search, Stochastic Field Simulation, etc.)
Targeted Random Sampling

Failure

Safe
Targeted Random Sampling

Random (iid) sampling
Targeted Random Sampling

Stratified Sampling/
Refined Stratified Sampling
The biggest problem with Monte Carlo methods is that they are wasteful
• Many simulations in regions that are, statistically, of little or no use
• This is especially problematic for reliability analysis where probability of failure is very low
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Repeat until convergence criteria satisfied
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Repeat until convergence criteria satisfied
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Repeat until convergence criteria satisfied
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Notes:
- TRS constructs a piecewise approximation of the failure surface
- Need to identify at least one failure point per disjoint failure region
- Uses available information to inform stratification

Repeat until convergence criteria satisfied
Example 1: Initial Sample

Limit State:  \[ G(U) = -\frac{U_1}{4} + \sin(5U_1) + 4 - U_2; \quad U \sim N(0,1) \]

Probability of Failure:  \[ p_f \approx 4.154 \times 10^{-4} \]
Example 1: 200 Samples

Limit State: \[ G(U) = -\frac{U_1}{4} + \sin(5U_1) + 4 - U_2; \quad U \sim N(0,1) \]

Probability of Failure: \[ p_f \approx 4.154 \times 10^{-4} \]
Example 1: 500 Samples

Limit State: \( G(U) = -\frac{U_1}{4} + \sin(5U_1) + 4 - U_2; \quad U \sim N(0,1) \)

Probability of Failure: \( p_f \approx 4.154 \times 10^{-4} \)
Example 1: Comparison

Limit State: \[ G(U) = \frac{-U_1}{4} + \sin(5U_1) + 4 - U_2; \quad U \sim N(0,1) \]

Probability of Failure: \[ p_f \approx 4.154 \times 10^{-4} \]

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean ( P_F )</th>
<th>COV</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS ((10^8 \text{ samples}))</td>
<td>4.154e-4</td>
<td></td>
</tr>
<tr>
<td>FORM</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>SORM</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Importance Sampling (IS)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>SS((500; 50; 10))((~1000 \text{ samples}))</td>
<td>4.2641e-4 ((100))</td>
<td>66%</td>
</tr>
<tr>
<td>TRS ((9+491))((500 \text{ samples}))</td>
<td>3.9998e-4 ((100))</td>
<td>4%</td>
</tr>
</tbody>
</table>
Example 2: Comparison

Limit State*: $G(X) = \sum_{i=1}^{5} X_i - C; \quad X_i \sim \text{Exp}(\lambda_i); \quad \lambda_i = 1 \forall i; \quad C = 0.26715$

Probability of Failure: $p_f \approx 9.2300 \times 10^{-6}$

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean $P_F$</th>
<th>COV</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS (10^8 samples)</td>
<td>9.2300e-6</td>
<td>3%</td>
</tr>
<tr>
<td>FORM</td>
<td>1.3913e-4</td>
<td></td>
</tr>
<tr>
<td>SORM</td>
<td>1.4847e-4</td>
<td></td>
</tr>
<tr>
<td>IS (1000 samples)</td>
<td>8.3715e-6</td>
<td></td>
</tr>
<tr>
<td>SS(500; 50; 10) (~1000 samples)</td>
<td>9.8664e-6 (100)</td>
<td>50%</td>
</tr>
<tr>
<td>TRS (243+757) (1000 samples)</td>
<td>9.4830e-6 (100)</td>
<td>20%</td>
</tr>
</tbody>
</table>

Challenges & Future Directions

• Defining an appropriate initial stratification in high dimension
  – $3^{20} \sim 3.5 \times 10^9$
  – Use MCMC to explore the space and post-stratify

• Dynamic Problems & Stochastic Processes/Fields

• Other RSS designs
  – Variance Minimization for UQ
  – Sensitivity-based Sampling
  – Space-filling sample designs
  – Etc.