Uncertainty, Sample Size and the 95/95 Tolerance Limit

Mikko Jyrkama and Mahesh Pandey

NSERC-UNENE Industrial Research Chair
Department of Civil and Environmental Engineering
University of Waterloo, Waterloo, Ontario, Canada

June 4 - 7, 2017
Niagara Falls, Ontario
Uncertainty and the 95/95 Tolerance Limit

Outline

- Study objectives
- Methodology
  - Basic idea and example
  - Wilks’ method
- Tolerance limits
- Sample size determination
- Estimation vs. ensuring compliance
- Summary and conclusions
The objective of safety analysis is to demonstrate that some important or critical variable (e.g., pressure, temperature, etc.) meets a specified acceptance (i.e., regulatory, safety, etc.) limit.

Monte Carlo simulation is often used to derive the distribution of the critical variable:
- i.e., probabilistic assessment
- Hence, the limit must be met with high probability.

**Key Question:** How many simulations should be performed to ensure sufficient confidence in the results?
- Depends on what quantity is being estimated.
Study Objectives

The objective of this study is to illustrate how an upper bound value (e.g., 95th percentile) can be estimated with high confidence (e.g., the 95/95 value) using a very small number (e.g., 59) of simulation trials.

- Based on the original work by Wilks (1941)
- The estimation is independent of the underlying probability distribution (i.e., the unknown distribution of the model output from Monte Carlo simulation)

Key distinction is made between

- Estimation vs.
- Ensuring compliance
The uncertainty or confidence associated with upper (or lower) bound estimates was first rigorously derived by Wilks in 1941.

- Often referred to as the Wilks’ method (or formula).

The original derivation is quite tedious and mathematically intensive.

Our paper provides a simpler derivation based on the concept of order statistics and the Binomial distribution.

- Please refer to our paper in the conference proceedings for more details…
Basic Idea

- In a random sample of 100 values, the 95th percentile value can be estimated as the 95th highest value (in the sorted list)
  - Need to consider ordering, or order statistics
- Consider a random sample of only four different values \( x_1, x_2, x_3, x_4 \) from some continuous distribution
- Question: What is the probability that the third highest value (i.e., next to largest order statistic) will be less than, say 1?
  - At least 3 of the values must be less than 1
Note that the order of the sample values can change.

- The various combinations of $k$ out of $n$ values can be evaluated using the **Binomial distribution**.
The upper bound 100(1 – α)% confidence limit for any percentile \( \pi_p \) can be stated as

\[ P(Y_i > \pi_p) = 1 - \sum_{k=i}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = 1 - \alpha \]

Probability (or confidence) that a particular \( i^{th} \) order statistic (in a sample of size \( n \)) is greater than a particular \( p^{th} \) percentile value of the distribution.

**Unknowns**

- \( i \) is the \( i^{th} \) order statistic (i.e., \( i^{th} \) highest value)
- \( n \) is the sample size
- \( p \) is the percentile value
- \( \alpha \) is the “level of significance” ((1 – \( \alpha \)) is the confidence)
The one sided (upper) confidence bound is equivalent to the upper tolerance limit

\[ P(Y_i > \pi_p) = 1 - \sum_{k=i}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = 1 - \alpha \]

This equation can be used to back-calculate the sample size that would be required to make the probabilistic statement true at a given level of confidence (for a chosen order statistic)

- This derivation is independent of the underlying probability distribution
The upper confidence limit associated with the $n^{th}$ highest (i.e., maximum) value in a sample of $n$ is

$$P(Y_n > \pi_p) = 1 - \sum_{k=n}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = 1 - p^n = 1 - \alpha$$

Solving for the sample size $n$

$$n = \frac{\ln(\alpha)}{\ln(p)}$$

For the 95/95 tolerance limit, $\alpha$ is equal to 0.05 and $p$ is 0.95, resulting in $n$ equal to 58.4
Choosing the highest (or maximum) value in a random sample of 59 values (e.g., from Monte Carlo simulation) will satisfy the 95/95 probability/confidence criterion

- Identical statement can also be made for a lower bound value using the minimum value in the sample

In other words, 95 % of the time (i.e., 19 times out of 20), the maximum value in a sample of 59 will be greater than the “true” or actual 95th percentile value of the distribution

- 5 % of the time it will be lower
The sampling distribution of the maximum value in a sample of 59
Wilks’ formula guarantees that 95% of the time, the maximum value (in a group of 59) will be greater than the true 95th percentile value.

But, there is a high likelihood of getting a value greater than the 98th percentile.

May result in an upper bound estimate that is too high or unrealistic.

The estimation can be improved by performing a higher number of simulations, i.e., using Wilks’ method to find the required sample size for different (lower) order statistics.
Sampling distributions of the $m^{th}$ highest values in a given number of simulation trials.

<table>
<thead>
<tr>
<th>$m^{th}$ Highest Value</th>
<th>Number of Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2326</td>
</tr>
<tr>
<td>40</td>
<td>1008</td>
</tr>
<tr>
<td>20</td>
<td>548</td>
</tr>
<tr>
<td>10</td>
<td>311</td>
</tr>
<tr>
<td>4</td>
<td>153</td>
</tr>
<tr>
<td>3</td>
<td>124</td>
</tr>
<tr>
<td>2</td>
<td>93</td>
</tr>
<tr>
<td>1</td>
<td>59</td>
</tr>
</tbody>
</table>
Comments

- In all cases, as before, the estimated value will always be greater than the true 95th percentile with 95% probability.

- However, choosing a higher number of simulations (i.e., lower order statistic in the sample) generally results in a less conservative estimate of the upper bound value:
  - The estimated value is likely to be closer to the actual true 95th percentile value, rather than some higher percentile.
Ensuring Compliance

- Rather than estimation, Wilks’ method is highly useful for **demonstrating compliance** with a specified acceptance limit
  - e.g., 2200°F peak cladding temperature of a fuel rod
- The 95/95 rule is met if **none** of the 59 simulated values exceed the critical limit
  - The maximum value would represent the 95th percentile value (likely to be higher)
  - Since the maximum value is less than the acceptance limit, then 95 % of all possible values of the random variable (e.g., simulated peak cladding temperature) should by definition be less than the acceptance limit (this statement is true with 95 % probability)
Given the small sample size, this approach naturally raises the key issue of **model uncertainty**

- e.g., estimating the 95th percentile with 50% confidence requires only 14 simulations (95/50 rule)

The model output (i.e., the small number of sample values) depend entirely on the

- Distributions (i.e., probabilistic models) of the inputs
- Computational model itself

Need to perform comprehensive

- Model validation and verification
- Sensitivity and uncertainty analysis
Summary and Conclusions

- The origin of Wilks’ sample size formula was derived using simple examples
  - Please see actual paper for more details
- The resulting expression can be used to determine the required sample size for any probability and confidence level
  - e.g., 99/99 criterion would require 459 simulations
- Rather than estimation, the method is most useful for ensuring compliance with a particular acceptance limit