CISC-820 Project 3 Report

Image Adhikari, Patrick Philippy

November 2024

1 Conclusions

CI	Validity	Alpha	Asymptotic	Notes
1	Valid	0.00	Yes	$\alpha = 0$ is not very informative.
2	Invalid	N/A	N/A	Fails consistently for highly right-skewed distributions.
3	Valid	0.05	Yes	
4	Invalid	N/A	N/A	Fails consistently for distributions near 0 or 1
5	Valid	0.1	No	
6	Invalid	N/A	N/A	Fails consistently for distributions near 0
7	Valid	0.1	No	
8	Invalid	N/A	N/A	Fails consistently for highly left-skewed distributions.
9	Valid	0.01	Yes	
10	Valid	0.05	No	

For more detailed data and visualizations, please see our Google Drive for this project.

2 Methods

2.1 Data Design

To generate the dataset used in our experiments, we implemented a sample generator capable of producing data from different probability distributions while ensuring all values remain within the interval $[0, 1]$. The generator supports the following distributions:

- **Bernoulli Distribution**: Produces binary outcomes (0 or 1) with a specified probability of success p. For instance, setting $p = 0.5$ ensures an equal likelihood of generating 0 or 1.

- **Uniform Distribution**: Generates continuous values between a lower bound and an upper bound. For example, $lower = 0$ and $upper = 1$ create a uniform distribution over the range $[0, 1]$.

- **Noisy Distribution**: Samples were pulled from a $\mathcal{N}(0, 1)$ Gaussian distribution until $1,000,000$ samples fell in the interval $[0, 1]$. The mean of this dataset was computed empirically, and treated as the true mean for subsequent sub-samples.

For each distribution, data was sampled in matrices of size $n \times m$, where n represents the number of data points per experiment, and m represents the number of experiments. The true mean of each dataset was calculated analytically based on the distribution parameters in order to evaluate the confidence intervals.

The generated datasets formed the basis for analyzing the performance of ten functions designed to compute confidence intervals. We evaluated these functions under varying conditions of data size and distribution type to assess their validity and robustness.

For further details on data generation and metrics, please see our GitHub.

2.2 Procedure

We define the following to test the validity of the confidence interval (CI) methods:

- **Exact Confidence Interval**: If the fraction missed remains below some α across all sample sizes and distributions, the function is valid.

- **Asymptotic Confidence Interval**: If the fraction missed exceeds α for small N but converges to α for large N, the function may still be valid asymptotically. We consider convergence to be small deviation relative to α after sufficient N.

- **Invalid Confidence Interval**: If the function is not an exact CI or an asymptotic CI, it is considered invalid.

A confidence interval is considered **Exact** only if it meets the criteria for all tested distributions. A confidence interval is **Invalid** if it is found to be invalid for any dataset. We run our experiments with different values of mean for the Bernoulli distribution [0.1, 0.5, 0.9] and different values of lower and upper bounds for the uniform distribution ([0.0,0.3], [0.3,0.7], [0.7,1.0], [0.0,1.0]) to arrive at a conclusion.