Project 3: Quantitative Foundation

Garegin Grigoryan Nilesh Kumar

Introduction

In this project, we're given 10 functions that allegedly compute confidence intervals. The task is to identify correct functions and the confidence level for those correct functions.

Background

A confidence interval is a pair of random variables A_N, B_N, such as for a dataset of independent identically distributed data $X_1, X_2, \ldots, X_N$: $Pr[A_N <= \theta <= B_N]=1-\alpha$, where θ is a property of the distribution generating X_i, α is a confidence level, and *Pr* is the probability that demonstrates the reliability of the estimation procedure.

Validity Table:

C.I	**Valid**	**Confidence Level**	**Asymptotically Only**	**Experiment References**
1.	Valid	1	Yes	
2.	Invalid			Theta = 1.000000e-02 Miss rate approaches 100 for higher values of N
3.	Valid	0.93	Yes	
4.	Invalid			a= 0, b= 1.000000e-02 for all values of N
5.	Valid	≈ 0.9	No	
6.	Invalid			a= 0, b= 1.000000e-02 for all values of N
7.	Valid	≈ 0.88	No	
8.	Invalid			Theta = 1.000000e-02 miss rate becomes 100% with higher values of N
9.	Valid	≈ 0.97	Yes	Bernoulli experiments

10.	Valid	≈ 0.94	No	

* N = number of samples

Input Data Design:

We test different inputs by varying the number of samples and distributions.

- ## Bernoulli Distribution

  We create a vector of the form [0:0.01:1] and use each of its values as a probability of success in Bernoulli with N = [10,100,1000,10000,20000,30000]. We use the extra values of N (20000,30000) for functions 2 and 8 because for every N after 100 it seemed to give 100% miss rate for some experiments so we wanted to make this is true for higher values of N.

- ## Uniform Distribution

  We create a vector of the form [0:0.01:1], we assign each value of this vector to **a** except for the last one and for each value of **a** we assign **b** all values greater than **a** in the vector with N = [10,100,1000,10000].

- ## Normal Distribution

  We did not construct any systematic input space for normal distribution because it took us 14 hours to run all the experiments on a uniform. Also, from previous experiments, we realized that most of the invalid behaviors are at the samples close to 0 and 1 so we just cherry-picked mu and sigma values that gave samples closer to 0, 1 and in the middle.

Testing Procedure:

We run experiments for the above-mentioned inputs and once we find any function that does not satisfy the definition of the valid confidence interval for any distribution for all values of N (or approaches 100% miss with increasing N) we consider it is invalid. To calculate confidence level we are using the experiment from valid experiments (smaller N values are excluded for asymptotically valid functions). If there is a noticeable difference in the hit percentages across the distributions or even within the distributions we use the worst performing experiments. As in function 7 our decision is based on bernoulli experiments. In cases where the functions give an almost similar hit percentage across and within distributions we roughly use the average of the values as in function 5 (all miss percentages are around 10%).