

Confidence Intervals

Khudayar Farmanli
Politecnico Di Torino
Data Science and Engineering
Torino, Italy
s276310@studenti.polito.it

Abstract—In this homework we try to understand how Accuracy changes based on different Confidence levels

I. INPUT PARAMETERS

- X - output of stochastic process, generated uniformly between 0 and 10
- Number of experiments - constant number
- Confidence level - defined as 90, 95 and 99
- Seed - for initializing the random number generator, defined as 8111996

II. FORMULAS ADOPTED

- Sample mean

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- Sample standard deviation

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2}$$

- Margin of Error

$$\delta = z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

- Confidence Interval

$$I = [\hat{x} - \delta, \hat{x} + \delta]$$

- Relative Error

$$\varepsilon = \frac{\delta}{\hat{x}}$$

- Accuracy

$$accuracy = 1 - \varepsilon$$

In our particular case instead of Normal distribution we will use so-called Student's t distribution which is calculated with $n - 1$ degrees of freedom unlike its counterpart and used generally when $n < 30$ where n is sample size. With increasing number of sample size it becomes closer to Normal distribution.

III. PYTHON FUNCTIONS ADOPTED

For solving this laboratory I used libraries like numpy to generate uniformly distributed random values, then calculated mean and std with pre-defined functions and scipy.stats which is also used for statistical operations and it helped to calculate Confidence Interval.

IV. RESULTS

Below we see Confidence Intervals of 90, 95 and 99 percents and their Accuracies particularly.

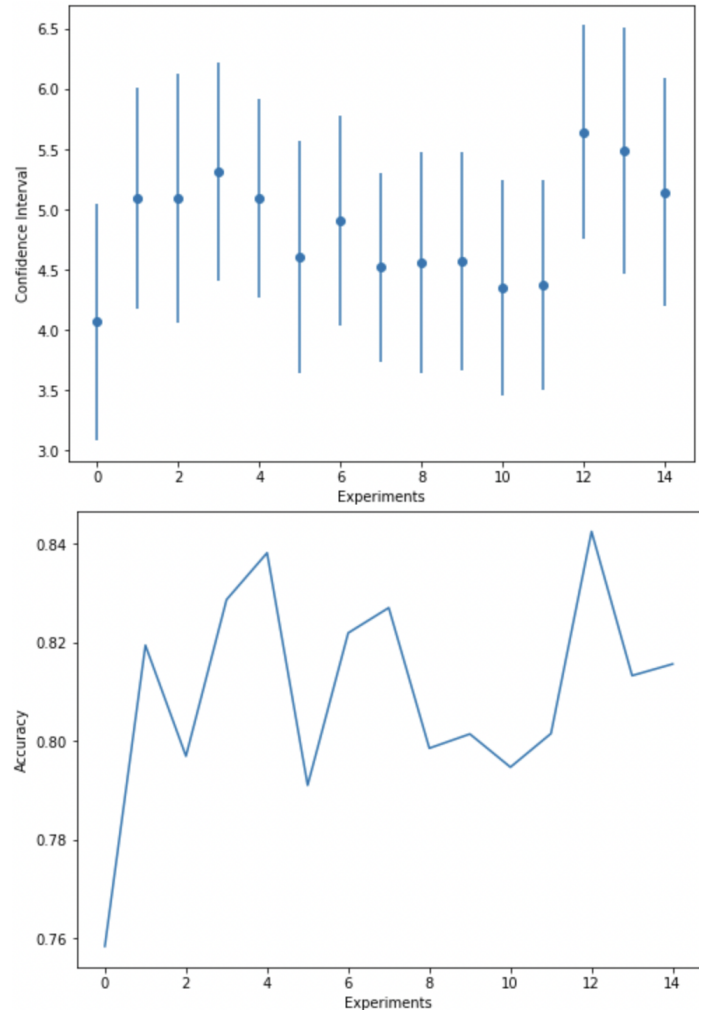


Fig. 1: Confidence Interval of 90 percent and its Accuracy

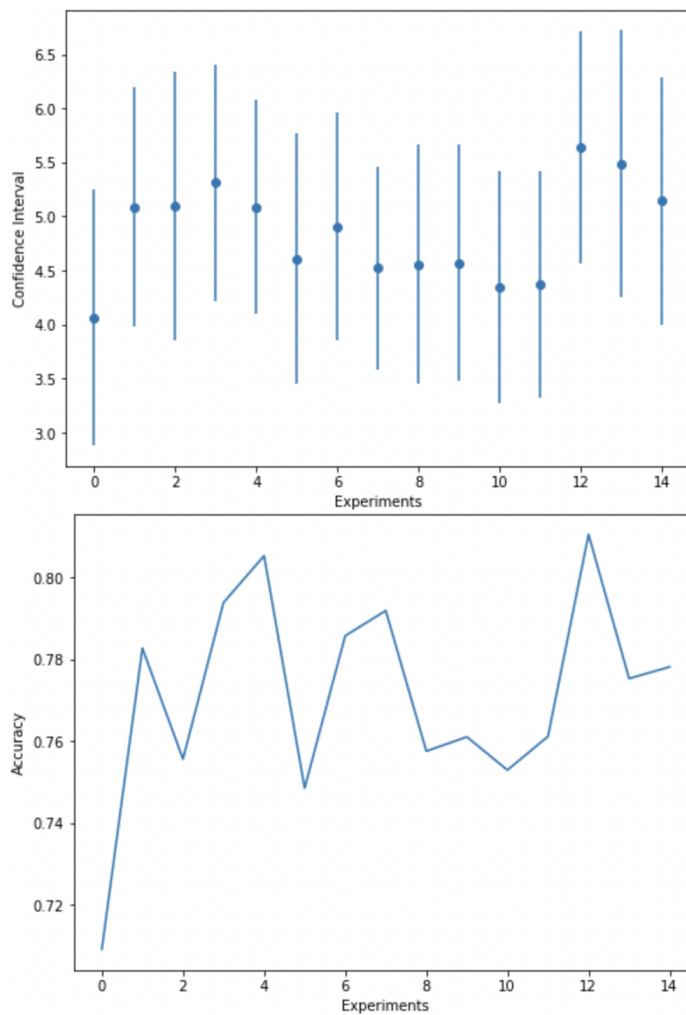


Fig. 2: Confidence Interval of 95 percent and its Accuracy

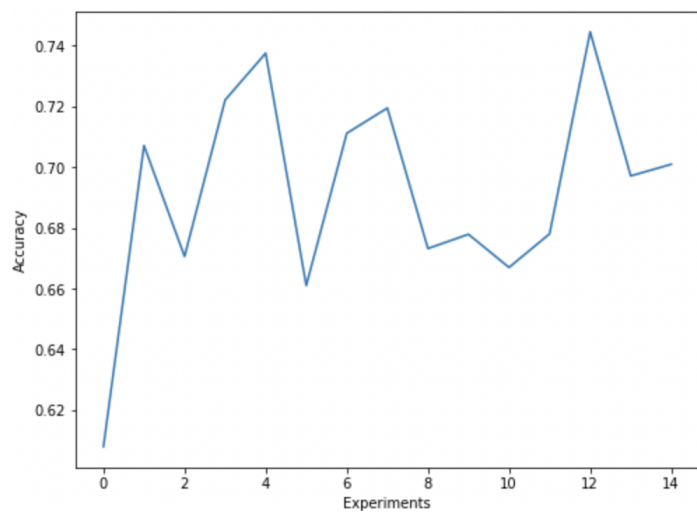


Fig. 3: Confidence Interval of 99 percent and its Accuracy

V. CONCLUSION

From the results we understand that increasing degree of Confidence level also widens our Confidence Interval and at the same time when we plot Accuracy of these confidence levels we see that the trade-off actually reveals here. Increasing number of confidence level corresponds to decreasing accuracy level.