

# Network analysis and simulation

## Homework 1

Davide Peron

### Exercise 1

In the first exercise two datasets are given and on those, a bunch of figures have been plotted, in which the data are showed and different measures of confidence are calculated on them.

The first dataset is a collection of execution times resulting from a first version of a program and a second more optimized one. In the follow the said figures are reported.

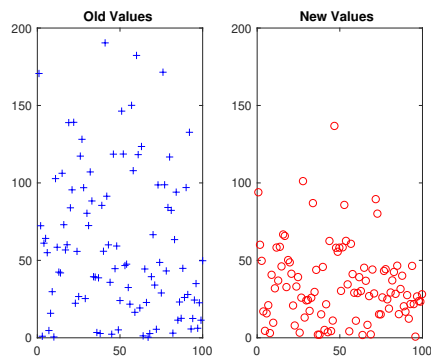


Figure 1: Plot of the data

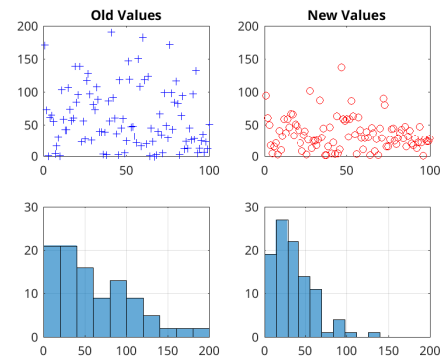


Figure 2: Data plotted also in histograms divided in 10 bins (Figure 2.1)

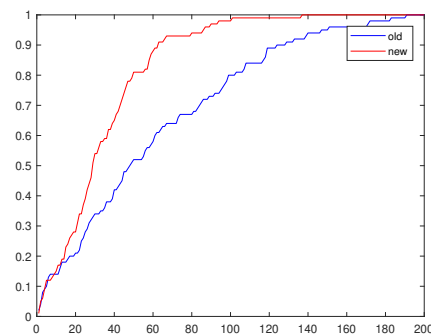


Figure 3: Empirical distribution function of the data (Figure 2.2)

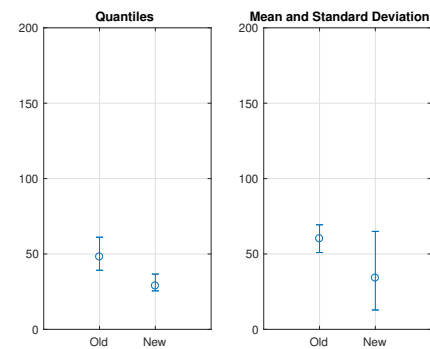
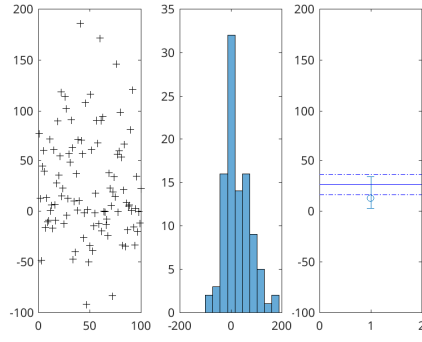


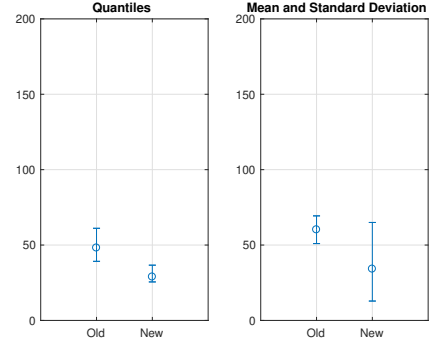
Figure 4: Box Plots of the data with Confidence Interval (CI) for median and mean (Figure 2.3)

### Exercise 2

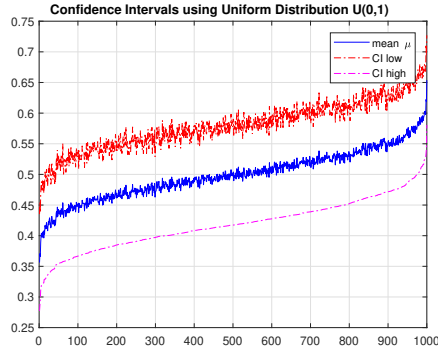
Executing the script correspondent to the second exercise, we found that 56 experiments the CI does not contain the true value of the mean.



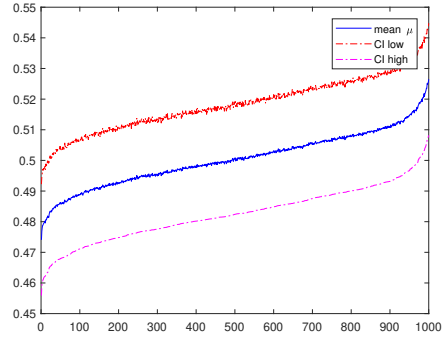
**Figure 5:** Difference between old and new data (Figure 2.7)



**Figure 6:** Box Plots of the data with CI for median and mean (Figure 2.3)



**Figure 7:** Results of the experiment with  $n = 48$



**Figure 8:** Results of the experiment with  $n = 1000$

In Figure 7 and Figure 8 are reported the value of the sample mean and its CI (with  $\gamma = 0.95$ ) for each experiment, sorted based on the lower extreme of the CI and using a different number of random variables in each experiment. Note how, increasing the number of random variables per experiment, the mean width of the CI get lower. Furthermore, the width of the CI is not constant.