Research Internship
How Many Growth Curves are Enough?

Topic profile

math

coding

biology

Tags

#statistics

#experimental design

#theoretical foundations

#interdisciplinary research

Supervision

Matthias Fuegger

CNRS Researcher at ENS Paris-Saclay

Gabriel Le Bouder

Postdoc at ENS Paris-Saclay

Thomas Nowak

Professor at ENS Paris-Saclay

What's wrong with n = 3?

Current practice in microbiology is to perform each experiment three times. Repeating the same experiment has the double purpose of being able to eliminate singular outliers and of having more data to increase the statistical robustness of the findings. The choice of $n = 3$ replicates, however, does not have a solid theoretical backing. This is problematic for two reasons: Firstly, too few samples lead to imprecise estimation of the system's parameters. Secondly, many statistical tests, like the t-test, rely on the fact that the empirical average is normally distributed. While this is asymptotically true for $n \to \infty$ by the Central Limit Theorem, it is unclear how far from normal the distribution for $n = 3$ is, and what this means for the validity of the test.

What we are looking for

We value a curious and driven attitude. An ideal candidate is inclined to statistics, probability theory, microbiology, and coding (in Python).

Research

The goal of this internship is to explore the validity of statistical tests on microbiological data. As a prototypical example, we will use growth-curve measurements. A possible way of tackling the problem is:

- Starting from a kinetic model of cell duplication of bacteria, create a synthetic data set of cell-density measurements over time.
- Compare the normality of the empirical averages for different numbers n of experimental replicates and assess the impact of non-normality on the validity of various statistical tests.
- Propose a new theoretically founded value for n and give an estimate of the error made in the statistical tests, or, failing that, develop new tests for the specific case of growth-curve data.

The team

You will be part of an interdisciplinary research team at ENS Paris-Saclay near Paris, working on different aspects of synthetic biology, distributed computing, and circuit design.

You are interested or would like to join us?

Please send us your questions or, in case you would like to apply, a short statement of interest and a curriculum vitae, to Matthias Fuegger (mfuegger@lmf.cnrs.fr) and Thomas Nowak (thomas@thomasnowak.net). Applications received until December 15, 2023 will receive full consideration. The start date of the internship is flexible, but the goal is to start in spring or summer 2024.