**Statistics for biomedical scientists**

**21- 23rd October 2019**

While looking at the experimental data we try to recognize patterns and to make sense of the observations. However, our intuition is often wrong, and so there is a need to impose some objectivity and the methods by which observations are converted into knowledge. Biostatistics is the branch of statistics that provides such methods. Understanding, interpreting data, drawing the conclusions are linked to comprehension of a prescribed set of mathematical expressions.

This course will introduce participants to the basics of statistics on a more intuitive way without being encumbered by equations. This introductory course is to foster an appreciation for the role of statistics and associated data analysis approaches in research and in our everyday lives.

**Aim of the course:**

To provide an introduction to biostatistics by explaining statistical principles with a focus on the scientific interpretation of statistical tests rather than on the mathematical logic of the tests themselves.

**By the end of the course, the participants should be able to:**

- Describe the main terms of descriptive statistics

- Explain the p-value and its meaning

- Explain the problem of multiple testing

- Discuss the tests to account for multiple testing

- Perform the main statistical test on a data set

- Describe the Bayesian statistics

**Target group:** PhD candidates in the beginning of their PhD trajectory

**Prerequisites:** The basics knowledge of statistics, probability theory and combinatorics, basics knowledge of one of the programming languages (preferably Perl)

**Duration of the course:** 3 days

**Location:** GIGA B34 +5

**Workload:** 3 days x 8 hours per day = 24 hours

**Educators:** Prof. Michel Georges (GIGA, ULiege)

**Course Syllabus/schedule**

**Day 1: Statistical significance. 9:00-18:00. 13:00-14:00 - Lunch**

* Definition of p-value and confidence intervals
* Permutation tests
* Bootstrapping
* Accounting for confounders

**Day 2: Multiple testing. 9:00-18:00. 13:00-14:00 - Lunch**

* Multiple testing: the issue.
* Adjusting the thresholds
* Exploiting the p-value distribution: false discovery rate

**Day 3: Advanced topics. 9:00-18:00. 13:00-14:00 - Lunch**

* Likelihood based tests
* Bayesian statistics