# Quantitative Understanding in Biology I
# Q1 2025
# Course Syllabus

Tuesday, August 26th, 2025 – Thursday, October 9th, 2025
Tuesdays and Thursdays, 5:30 PM – 7:30 PM
Location: LC-504
Course Directors:


- Jason Banfelder (✉ jbanfelder@rockefeller.edu)
- Luce Skrabanek (✉ las2017@med.cornell.edu)


This course will be fully graded: (Honors, High Pass, Low Pass, Fail)
Course web site: `http://physiology.med.cornell.edu/people/banfelder/qbio`

Teaching assistant:


- Aria Dandawate (✉ ard4008@med.cornell.edu)


Teaching assistant duties:

- Conducting review sessions of classroom lectures
- Conducting review sessions for exams
- Conducting one-on-one tutoring sessions
- Being available to students by holding regular office hours
- Grading homework

## Summary

This course will prepare students to apply quantitative techniques to the analysis of experimental data. To emphasize both practical and theoretical skills, the completion of several projects will be required. Students will be well positioned to meet the emerging requirements of funding agencies for formally planned experiments and fully reproducible and documented data analysis methods.

Specific topics include: practical aspects of data formatting and management; graphical, mathematical and verbal communication of quantitative concepts; a review of statistics, with emphasis on the selection of appropriate statistical tests, the use of modern software packages, the interpretation of results, and the design of experiments; the formulation, evaluation, and analysis of mathematical models of biological function, with an emphasis on linear and non-linear regression, determination of model parameters, and the critical comparison of alternative models with regard to over-parameterization.

## Class Sessions

1. Quantifying a Sample Distribution
   Tuesday, 26 August, 2025

   summary statistics, measurement variation vs. biological variation

2. Introduction to R Part I
   Thursday, 28 August, 2025

3. Probability Density Functions and the Normal Distribution
   Tuesday, 2 September, 2025

   binomial, Poisson, and normal distributions, testing for normality: an introduction to formal statistical tests, qqplots vs. formal tests

4. Introduction to R Part II
   Thursday 4 September, 2025

5. Confidence Intervals and Contingency Tables
   Tuesday, 9 September, 2025

   t-tests, working with proportional data, why CIs are more informative that p-values, study types (retrospective, prospective, and cross sectional), working with rare events

6. p-Values and Formal Statistical Testing
   Thursday, 11 September, 2025

duality between p-values and CIs, statistical vs. biological significance, choosing an appropriate $\alpha$, type I and type II errors

Problem Set #1: Understanding Type I error rates

7. Statistical Power and Experimental Design
   Tuesday, 16 September, 2025

   why you cannot just add a few N to your dataset when $p > 0.05$, appreciating the economics (in time and money) of experimental design, statistical vs. biological significance revisited, designing experiments around hard-to-obtain samples

8. Multiple Hypothesis Testing and Non-parametric Tests
   Thursday, 18 September, 2025

   from Bonferroni to False Discovery Rate, dealing with non-normal data, t-tests revisited

9. Problem Set #2: the optimal stopping problem
   Tuesday, 23 September, 2025

   the role of simulation in statistics, non-parametric tests

10. Bayesian Methods
    Thursday, 25 September, 2025

    how to incorporate prior knowledge into statistical models

11. Correlation vs. Linear Regression
    Tuesday, 30 September, 2025

    introduction to modeling in R; why $r^2$ is not the whole story

12. Fitting Model Parameters to Data
    Thursday, 2 October, 2025

    non-linear regression, a statistical view of curve fitting, confidence intervals revisited

13. Quantitative Comparison of Models and ANOVA; Problem Set #3
    Tuesday, 7 October, 2025

    how to avoid over-fitting, F-test and AICs, F-tests as a means of parameter estimation

14. Graphics with ggplot
    Thursday, 9 October, 2025

## Books and Materials

Students will need a laptop computer on which they can install software (R[1] and R Studio[2]), and bring to class. Both packages are free, and run on recent versions of Linux, Mac OS X, and Microsoft Windows.

While the course does not require the use of a specific textbook, the following resources are recommended.

- *Intuitive Biostatistics*, Harvey Motulsky

  One of the most accessible introductions to statistics.

- *The Art of R Programming*, Norman Matloff

  One of the more comprehensive introductions to R.

- *R for Everyone*, Jared Lander

  Less in-depth than the above, but covers both basic use of R and basic statistics in a single, accessible text.

- *Practical Computing for Biologists*, Haddock and Dunn

  Covers many computing topics not covered in this class. Recommended for students considering a computational lab for a rotation or thesis.

- *R for Data Science*, Grolemund and Wickham

  Introduction to the tidyverse, focusing on importing, wrangling, exploring, and modeling your data and communicating the results. Some experience with R advised. Also available online at `http://r4ds.had.co.nz/`.

## Assessment

This class will be graded according to the usual WCGS scale (Honors, High Pass, Low Pass, Fail).

Grades will be determined based on several take-home problem sets.

All students will also be asked to complete a survey at the end of the semester, soliciting feedback on the course to inform its content and format in future years.

[1] `https://www.r-project.org/`
[2] `https://www.rstudio.com/`