# Quantitative Neuroscience Core

| | | |
|---|---|---|
| *Instructor* | Joshua Gold | jigold@pennmedicine.upenn.edu |
| *TA* | Andy Revell | Andre.Revell@pennmedicine.upenn.edu |
| *Required meeting times* | MWF | 9–10 am |
| *Optional discussion section/office hours* | TBD | |
| *Meeting location* | Barchi Library | |

## Introduction

This course is designed to be an overview of quantitative approaches used for rigorous and reproducible neuroscience research. This course does not cover statistics in a traditional way, in the sense that we will not provide a comprehensive survey of statistical tests, nor will we dive very deeply into formal mathematical derivations of those tests (information about such things can be found in textbooks and all over the web). Instead, we will focus on teaching you to apply quantitative approaches to your thinking about neuroscience research from beginning to end, including defining clear hypotheses; designing experiments to test those hypotheses; collecting, visualizing, analyzing, and interpreting data in reference to those hypotheses; and keeping effective and transparent records at each stage to ensure rigor and reproducibility.

There are two main components to the course. The first component consists of a series of four modules, each of which is designed to use a specific example from neuroscience to illustrate a set of quantitative approaches and tools. The second component consists of group projects that focus on designing and implementing quantitative analyses for existing data sets (e.g., from your rotation project).

## Learning Objectives

**1) Develop good habits for transparent, reproducible science**. Transparency is the idea that none of your data or methods should be hidden. Reproducibility is the idea that you should be designing, conducting, and analyzing experiments in a way that maximizes the probability that someone else doing the same experiments would come to the same conclusions. To support these ideas, we will incorporate into the course the use of several on-line tools that, even if you do not end up using these particular tools in your own research, will help establish good habits for record keeping (we will use LabArchives electronic notebooks, https://researchnotebooks.upenn.edu), version control for code (we will use GitHub, https://github.com), and data storage (we will use PennBox: https://upenn.app.box.com).

**2) Learn to think about statistics in the context of good experimental design**. The question “what statistical test should I use?” can be answered only after answering more basic questions, like “what are the alternative hypotheses that I am testing?”and “how well does my experimental design allow me to distinguish those hypotheses?”

**3) Learn foundations of statistical reasoning, particularly how to think about randomness using probability distributions**. Even the most sophisticated statistical procedures are ultimately about distinguishing signal from noise. This ability depends on understanding what is meant by “noise”, or randomness. The primary mathematical tool for quantifying and manipulating randomness is the probability distribution, which describes the probability of obtaining all possible values of a quantity of interest (e.g., the outcome of an experiment). We therefore will spend some time learning about probability distributions and then build on those concepts to better understand how to use probability distributions to make inferences.

**4) Learn to visualize your data effectively to lay bare your statistical reasoning**. Ultimately your ability to convince other people that you have a robust finding will not depend on the results of a statistical test but rather on your ability to show the finding in a clear and compelling way; that is, in a way that is transparent in terms of what you measured, clearly reflects the experimental design, and illustrates both the signal and noise that you found. We will focus on specific ways to visualize data effectively throughout the course.

## Course Resources

Most course resources will be in two places:

1) Discussions listed on this page.

2) Exercises, tutorials, and sample code on the NGG GitHub Statistics Repository.

## Using Matlab

We will use Matlab (https://www.mathworks.com) in this course, so it will benefit you to have at least a rudimentary understanding of how to use it. It is available for free to all BGS students (please contact Christine for instructions on how to get it). We suggest that you get a copy as soon as possible and learn to use its basic functionality.

### How to use the tutorials and exercises

Numerous class sessions will involve in-class discussions and homework involving Matlab. For Matlab-based tutorials (e.g., this one that we will cover early in the course), you should download the code from GitHub to your computer, then go through the tutorial line-by-line, executing one line of code at a time. The goal of these tutorials is to give you a detailed perspective on a particular topic, and how to implement various concepts in Matlab code. For

Matlab-based exercises (e.g., the "Quantal release" exercises here that we will cover early in the course), you should try to answer the questions yourself in Matlab; answers are given in posted files on GitHub that you can then use to check your answers.

**Resources for Learning Matlab**

- From Mathworks: https://www.mathworks.com/help/matlab/getting-started-with-matlab.html
- Coursera: https://www.coursera.org/learn/matlab
- Wallisch et al, *Matlab for Neuroscientists* (https://www.sciencedirect.com/book/9780123838360/matlab-for-neuroscientists)
- The summer Matlab course offered by the NGG

*Note for students who prefer Python: the long-term plan is to teach the course using Python, but we are not there yet. If you prefer to use Python now, please come talk to me – I am all for it but want to work out the details on an individual basis.*

## Other External Resources

- The Society for Neuroscience, Promoting Awareness and Knowledge to Enhance Scientific Rigor in Neuroscience
- The Journal of Physiology, *Statistical Reporting Guidelines*
- BGS guidelines on the Responsible Conduct of Research (RCR) and Scientific Rigor and Reproducibility (SRR)
- Motulsky, H. *Intuitive Biostistatistics*
- Collected readings on quantitative rigor
- Tutorials and answers to exercises on the NGG statistics GitHub repository

## Grading

Grades are based on: 1) class participation, including engagement in discussions (20%); and 2) a final project involving three in-class presentations (20% each) and electronic records of analysis strategies and code (20%).

For our philosophy of grading, see here.

## PART 1: FOUNDATIONS

| | | | |
|---|---|---|---|
| **Wed** | **1-Sep** | **Introduction I: Overview and Goals** | **Please read before class and be prepared to discuss in class:**<br>Platt, J.R. (1964) Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. Science146, 347-353.<br>Kass, R.E. (2011) Statistical Inference: The Big Picture. Statistical Science 26(1). |
| **Fri** | **3-Sep** | **Introduction II: Reproducibility and Transparency (Revell)** | **Read and be prepared to discuss:**<br>Record Keeping: Laboratory Notebooks<br>Record Keeping: Algorithms<br>Record Keeping: Data<br>**Sign up for the following accounts (if you haven't already):**<br>LabArchives (through Penn)<br>GitHub (Links to an external site.)<br>PennBox (through Penn) |
| **Mon** | **6-Sep** | **LABOR DAY -- NO CLASS** | |
| **Wed** | **8-Sep** | **Introduction III: Frequentist versus Bayesian Approaches** | **Complete Colab tutorial:**<br>Frequentist versus Bayesian approaches |
| **Fri** | **10-Sep** | **Data Visualization I: Principles (Revell)** | |
| **Mon** | **13-Sep** | **Data Visualization II: Examples (Revell)** | **Find a figure/graph from a paper you think displays the distribution of their data well or poorly. Post it in the Canvas course discussion.** |
| **Wed** | **15-Sep** | **Probability Distributions I: Concepts** | **Complete and be prepared to discuss these Colab tutorials:**<br>Samples and Populations<br>Probability Distributions Overview<br>Bernoulli Distribution<br>Binomial Distribution<br>Exponential Distribution<br>Gaussian (Normal) Distribution |

| | | | |
|---|---|---|---|
| | | | Poisson Distribution<br>Student's t Distribution |
| **Fri** | **17-Sep** | **Probability Distributions II: Binomial Distribution Case Study** | **Answer questions from the Neuroscience Example ("Quantal release") case study in the Binomial distribution Colab tutorial and post your answer code on GitHub**<br>Binomial Distribution |
| **Mon** | **20-Sep** | **Probability Distributions III: Confidence Intervals and Bootstrapping** | **Complete and be prepared to discuss this Colab tutorial:**<br>Confidence Intervals and Bootstrapping |
| **Wed** | **22-Sep** | **Two-Sample Inference I: Experimental Design and Power Analysis** | **Read and be prepared to discuss:**<br>Button et al (2013), Power failure: why small sample size undermines the reliability of neuroscience<br>**Complete and be prepared to discuss this Colab tutorial:**<br>Error Types, P-Values, False-Positive Risk, and Power Analysis |
| **Fri** | **24-Sep** | **Two-Sample Inference II: Parametric Tests and Multiple Comparisons** | **Complete and be prepared to discuss these Colab tutorials:**<br>t-tests<br>Multiple comparisons |
| **Mon** | **27-Sep** | **Two-Sample Inference III: Nonparametric Tests** | **Complete and be prepared to discuss this Colab tutorial:**<br>Simple Non-Parametric Tests |
| **Wed** | **29-Sep** | **Measures of Association I: Correlation** | **Complete and be prepared to discuss these Colab tutorials:**<br>Measures of association<br>Parametric correlation coefficient (Complete the exercises and post your answer code on GitHub)<br>Nonparametric correlation coefficient |

| | | | |
|---|---|---|---|
| Fri | 1-Oct | Measures of Association II: Simple Linear Regression | Complete and be prepared to discuss these Colab tutorials:<br>Measures of association<br>Simple linear regression (Complete the exercises and post your answer code on GitHub) |

| | | | |
|---|---|---|---|
| Mon | 4-Oct | Measures of Association IV: Nonparametric Correlation Case Study | Read and be prepared to discuss:<br>Aston-Jones and Cohen (2005), Figure 7<br>Joshi et al (2016), Figure 3<br>Optional: Review the code in the NGG GitHub Repository under "Examples/LC-Pupil/" that was used to generate Fig. 3 of Joshi et al.<br>Repository link |

| | | | |
|---|---|---|---|
| Wed | 6-Oct | QNC Modeling I: LATER Model Case Study | Read and be prepared to discuss:<br>Noorani (2014)<br>Some more readings just for fun:<br>RT at Penn I<br>RT at Penn II<br>RT at Penn III |

| | | | |
|---|---|---|---|
| Fri | 8-Oct | QNC Modeling II: RT Data Visualization | Run the Matlab tutorials in the NGG GitHub Repository under "Examples/LATER model/laterTutorial_plot*"<br>Repository link |

| | | | |
|---|---|---|---|
| Mon | 11-Oct | QNC Modeling III: Model Fitting | Run the Matlab tutorials in the NGG GitHub Repository under "Examples/LATER model/laterTutorial_model*"<br>Repository link |

**PART 2: APPLICATIONS (STUDENT PRESENTATIONS)**

| | | | |
|---|---|---|---|
| Wed | 13-Oct | PRESENTATION 1: HYPOTHESES AND | |
| Fri | 15-Oct | | |
| Mon | 18-Oct | | |
| Wed | 20-Oct | | |

| | | |
|---|---|---|
| Fri | 22-Oct | |
| Mon | 25-Oct | |
| Wed | 27-Oct | |
| Fri | 29-Oct | |

| | | |
|---|---|---|
| Mon | 1-Nov | PRESENTATION 2: DATA VISUALIZATION |
| Wed | 3-Nov | |
| Fri | 5-Nov | |
| Mon | 8-Nov | |
| Wed | 10-Nov | |
| Fri | 12-Nov | |
| Mon | 15-Nov | |
| Wed | 17-Nov | |

| | | |
|---|---|---|
| Fri | 19-Nov | PRESENTATION 3: HYPOTHESIS TESTING |
| Mon | 22-Nov | |
| Wed | 24-Nov | |
| Fri | 26-Nov | THANKSGIVING -- NO CLASS |
| Mon | 29-Nov | |
| Wed | 1-Dec | |
| Fri | 3-Dec | |
| Mon | 6-Dec | |
| Wed | 8-Dec | |