

Introduction to Statistics

MOL518, Lecture #8

Overview of the “basic” statistics unit

Today: Introduction to statistics and statistical functions in Python

Friday 2/13: What is Probability? What are probability distributions?
Introduction to the 4 “Great Distributions” in Biology

Monday 2/16: How do we simulate random numbers and distributions?

Wednesday 2/18: How do we compare two distributions? Introduction to statistical tests (and p-values)

Friday 2/20: How do we compare two variables? Introduction to correlation and linear regression

After that, we will have a unit on “intermediate” statistics

Monday 2/23: How do we fit curves to data? (beyond linear)

Wednesday 2/25: Review session with the TAs

Friday 2/27: Binary classification (setting thresholds)

Monday 3/2: How do we compare high-dimensional data? Clustering

Wednesday 3/4: How do we reduce the number of dimensions in a dataset using Principal Component Analysis (PCA)

Friday 3/6: Midterm exam

Are there any questions?

Now, over to Colab for the rest of the lecture!