

## Introduction to Biostatistics

### COURSE DURATION

This is an on-line, distance learning course and material will be available from:  
June 1 – 30, 2021

### INSTRUCTORS:

Martina Pavlicova, PhD  
Associate Professor of Biostatistics at  
CUMC  
722 W 168th St. 6th floor, room 635  
pavlicov@gmail.com  
212-305-9405

Sheila Nemeth, PhD  
[smv2123@columbia.edu](mailto:smv2123@columbia.edu)

### COURSE DESCRIPTION

The course will give an overview of the basic tools for the collection, analysis, and presentation of data in all areas of public health. The topics covered will include descriptive statistics; hypothesis testing; methods for comparison of discrete and continuous data including ANOVA, t-test, correlation, chi-squared analysis, linear and logistic regression, and non-parametric approaches. This course will provide a foundation for the skill-building courses and focused epidemiologic courses.

### PREREQUISITES

None.

### COURSE LEARNING OBJECTIVES

By the end of the course, participants will be able to:

- Apply numerical and tabular techniques commonly used to characterize and summarize public health data
- Describe basic principles of key concepts including hypothesis testing, type I and type II errors, power, and confidence bounds
- Identify appropriate statistical methods (ANOVA, correlation, linear regression, logistic regression) to be applied in a given research setting, apply these methods, and acknowledge the limitations of those methods
- Evaluate computer output containing statistical procedures and interpret it. Apply learned methods to other courses in public health

## COURSE READINGS

There are no required readings but the following textbooks are suggested if needed:

- “Principles of Biostatistics” by Marcello Pagano and Kimberlee Gauvreau, 2nd edition, Brooks/Cole, 2010.
- Less math oriented: "Fundamentals of Epidemiology and Biostatistics: Combining the basics" by Ray M. Merrill, 1st Edition, Jones & Bartlett Learning, 2013.

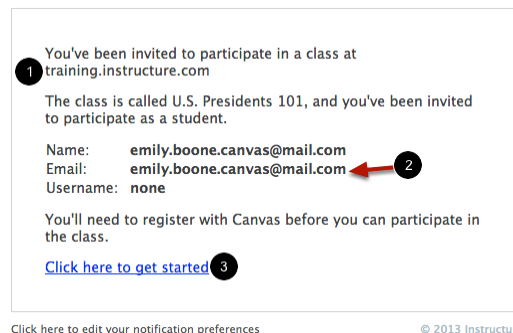
Additional material for the course will be provided to the students as needed.

## COURSE STRUCTURE

### CLASS WEBSITE

The class website will be hosted on Canvas

You will receive an invitation to join the course via email.

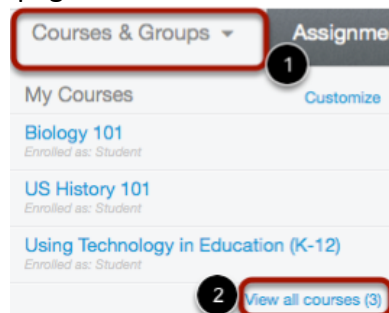


Once you receive the invitation, you can click on the “Click here to get started” link. You will need to create your own account (username and password) on Canvas

Once you have accepted the course invitation and created an account, you can always log into Canvas at: <https://canvas.instructure.com>

To access the course:

Select “Courses” at the top of the page.



Under “My Courses” select “Introduction to Biostatistics 2014”



For more help with Canvas, you can access the Student Quick Start Guide:  
<http://guides.instructure.com/m/8470>

The course website has 6 tabs on the left hand side of the page that are used to primarily navigate through the site.

1. The “Home” tab contains our introduction video for the course
2. A copy of this syllabus can be found of the “Syllabus” tab.
3. The course will primarily be set up in modules located in the “Modules” tab.
  - Each module will have videos of small lecture segments and review problems.
  - There will be approximately 3 modules per week.
  - All course material for the week will be posted on Sunday.
4. PDFs of lectures slides, review problems, and distribution tables can be found in the “Files” tab.
5. Office hours for both the instructor and the TA are held using the “Conferences” tab.
  - You will receive an email notification saying that you are invited to participate in a conference.
  - Once on the site, you can click on the ongoing conference and select “join conference.”
  - We recommend that you use headphones with a built-in microphone (e.g. iPhone headphones) rather than the built in computer microphone to reduce background noise.
  - We will give preference in office hours to questions that focus on the material covered in the current week rather than past or future material
6. The “Discussion” tab is an open forum where students enrolled in the course can communicate with one another.

## COURSE SCHEDULE

**Week 1 – Introduction to Biostatistics, Probability, Distributions, and Central Limit Theorem**

**Module 1** Learning Objectives: Introduction to Biostatistics; Study design; types of data; and Graphical and numerical representation of data.

<b>Module 2</b>	<u>Learning Objectives:</u> Introduction to probability; Law of total probability; and Bayes Theorem.
<b>Module 3</b>	<u>Learning Objectives:</u> Binomial distribution; Poisson distribution; Normal distribution; and the Central limit theorem.

<b>Week 2 – Introduction to Inference, Hypothesis Testing</b>	
<b>Module 1</b>	<u>Learning Objectives:</u> Confidence intervals for means and proportions; the t distribution; and Meaning of the confidence level.
<b>Module 2</b>	<u>Learning Objectives:</u> Introduction to hypothesis testing; Hypothesis testing in one sample for means and proportions; and Errors in hypothesis testing.
<b>Module 3</b>	<u>Learning Objectives:</u> Hypothesis testing for two sample proportions; and Hypothesis testing with two sample means.

<b>Week 3 – Power calculations and ANOVA</b>	
<b>Module 1</b>	<u>Learning Objectives:</u> Power calculations; and Relationships between power, effect size, and sample size.
<b>Module 2</b>	<u>Learning Objectives:</u> Introduction to analysis of variance (ANOVA); the F-distribution; and Bonferroni Correction.
<b>Module 3</b>	<u>Learning Objectives:</u> One-way ANOVA; two-way ANOVA; and two-way ANOVA with interaction.

<b>Week 4 – Linear Regression, Logistic Regression, and Non-Parametric Approaches</b>	
<b>Module 1</b>	<u>Learning Objectives:</u> Correlations; Simple linear regression; and Multiple regression.
<b>Module 2</b>	<u>Learning Objectives:</u> Odds ratios; Logistic regression; Chi-square test for independence; and Multiple logistic regression.
<b>Module 3</b>	<u>Learning Objectives:</u> Sign test; Wilcoxon sum-rank test for small samples; and Wilcoxon sum-rank test for large samples.